OPTIMIZING DECISIONS IN PRENATAL SCREENING FOR DOWN SYNDROME AND CAPACITY ALLOCATION IN A SCHOOL-BASED ASTHMA CARE MODEL

A Thesis
Presented to
The Academic Faculty

by

Jia Yan

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Industrial and Systems Engineering

Georgia Institute of Technology
December 2017

OPTIMIZING DECISIONS IN PRENATAL SCREENING FOR DOWN SYNDROME AND CAPACITY ALLOCATION IN A SCHOOL-BASED ASTHMA CARE MODEL

Approved by:

Professor Pinar Keskinocak, Advisor
School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor Turgay Ayer, Advisor
School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor David Goldsman
School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor Julie Swann
School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor Robert L. Schmidt
Department of Pathology and ARUP Laboratories
University of Utah School of Medicine

Date Approved: 22 August 2017

ACKNOWLEDGEMENTS

I would like to thank Drs. Turgay Ayer, David Goldsman, Pinar Keskinocak, Robert Schmidt, Julie Swann, for their time and efforts in serving as my committee members. I am especially grateful to my thesis advisors, Pinar Keskinocak and Turgay Ayer, for their great patience, guidance, encouragement, and support during my journey in the doctoral program.

My research was supported by George Family Foundation Grant. In addition, I am especially thankful to Drs. James Bost and Robert Palmer at Children's Healthcare of Atlanta, for supporting me and providing me the invaluable opportunity to work in the Outcomes Center as a research assistant during my Ph.D. study. It very much enriched my experiences in health systems and inspired me to incorporate operations research methodologies in the school-based asthma project.

This thesis would not have been completed without the unconditional support of my family. I am very much thankful to my parents, Jiannong Yan and Liping Lai, for their endless love. My most sincere gratitude is for my husband, Weijun Ding, who always stands by me and encourages me. It is my fortune to have you in this great time of my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This thesis consists of three decision making topics in two healthcare applications using operations research methodologies.

The first application is prenatal screening for Down syndrome (DS) which is a common type of chromosomal abnormality. Prenatal screening for pregnant women, based on multiple serum and ultrasound markers, is non-invasive and used to assess the risk of having a DS baby. A woman with a risk higher than a predefined risk-cutoff value of prenatal screening, i.e., a positive screening result, typically undergoes an invasive diagnostic test, such as amniocentesis, as a follow-up procedure to confirm that her fetus is affected. The risk-cutoff value of DS prenatal screening plays an important role. On one hand, a lower risk-cutoff value elevates the risk of false positives, and on the other hand, a higher risk-cutoff value increases the risk of false negatives. In practice, a one-size-fits-all type of risk-cutoff value of 1 in 270 is usually used for prenatal-integrated screening for DS. The objective of this application is to determine the optimal risk-cutoff values from the individual and population's perspectives.

The first topic focuses on the individual perspective. Women considering DS screening usually face two major adverse outcomes: undetected DS live births due to false negatives, and euploid procedure-related fetal losses due to false positives. Evidence shows that women perceive the two outcomes very differently. However, no guidelines exist for setting an appropriate risk-cutoff value based on women's different preferences. We capture the relative preferences using a ratio of weights (penalties), and formulate an optimization model to minimize the weighted sum of the two adverse pregnancy outcomes.

The second topic is from the population's perspective. Although using an appropriate one-size-fits-all cutoff value can achieve a high overall detection rate, it usually also brings in undesirably high false positive rates in older ages. Therefore, we explicitly add an upper-bound constraint to every single age for fairness and maximize the overall detection rate with those constraints.

The solution methodologies in this application combine the techniques of integer programming and Monte Carlo simulations. We find the preference-sensitive and age-specific risk-cutoff values have the potential to improve pregnancy outcomes and patient satisfaction. Our framework for DS prenatal screening can shed some lights on the optimal decisions in similar settings with a risk-cutoff value to designate positive or negative results.

The second application focuses on a capacity allocation problem in a school-based asthma care model, which is faced by Children's Healthcare of Atlanta to improve the health outcomes of pediatric asthma patients within metro Atlanta. In particular, the objective is to maximize the effectiveness of a school-based asthma program by deciding dynamically: (i) which schools to visit, and (ii) if a school is visited, which group of patients in this school to schedule for a limited number of clinical appointment slots. For this purpose, we propose a finite-horizon dynamic programming model which combines a clinic disease model of childhood asthma with operational decisions. We formulate a mixed-integer programming (MIP) for solving this model and propose two computationally-efficient and competitive heuristic solutions based on this MIP formulation. We parameterize our proposed models using data of a local public school district (Atlanta Public Schools district) and quantify the percentages of capacity allocated between patients in the worst illness states (treatment-prioritized capacity allocation) and patients in the moderate illness states (prevention-prioritized centered capacity allocation). We establish the following findings: (1) as capacity increases and planning horizon extends, more patients will be able to end up in the best

illness states, (2) our index-based heuristic consistently has a small gap compared to the optimal solution and it usually allocates more capacity to sickest patients, while the optimal solution usually has a smarter balance of allocating capacity between treating the sickest patients and preventing moderately sick patients from deteriorating, (3) capacity allocation are widely spread out among the public schools and driven by larger sizes of asthma patients who are sicker. We further quantify the value of these decisions, which may help decision-makers in daily operations.

CHAPTER I

INTRODUCTION

In health systems, many applications to improve health outcomes boil down to optimization problems, which can be too complicated to solve for optimality manually. It is useful to develop rigorous mathematical models to assist decision making. This thesis presents three topics on decision making optimizations with operations research methodologies. The first two topics focus on optimizing decisions in prenatal screening for Down syndrome (DS). Among those topics, we consider the problem from both individual and populations' perspectives. The third topic solves a capacity allocation problem in a school-based asthma care model.

1.1 Preference-Sensitive Risk-Cutoff Values for Prenatal-Integrated Screening Test for Down Syndrome

For a pregnant woman considering prenatal screening for early detection of DS, there are at least two major adverse outcomes of interest: undetected DS live births and euploid procedure-related fetal losses. Although evidence highlights the heterogeneity in women's preferences about the two adverse outcomes, the risk-cutoff value of 1/270 is commonly used for recommending a diagnostic test in practice. This topic focuses on the individual perspective.

1.1.1 Objectives

Since the guidelines of setting risk-cutoff values based on preferences are not in place, we propose that preference-sensitive risk-cutoff values should be considered instead of one-size-fits-all type of risk-cutoffs. The objective of this study is to assess the

impact of womens preferences for different pregnancy outcomes on the optimal risk-cutoff values for integrated screening, which has a higher detection rate compared to other maternal serum and ultrasound screening tests.

1.1.2 Methodology

We build a Monte Carlo simulation model of 100,000 singleton second-trimester pregnancies to assess the probabilities of DS live births and euploid procedure-related fetal losses for various risk-cutoff values, since there exist no explicit formulas to capture the relationship between risk-cutoff values and the adverse outcomes. To capture how undesirable some women may view an undetected DS live birth relative to a euploid procedure-related fetal loss, we use a ratio $W_1 : W_2$ of weights (penalties) assigned to these two adverse pregnancy outcomes. We formulate the problem as an unconstrained optimization to minimize the weighted sum of the two adverse pregnancy outcomes, and determine the optimal risk-cutoff value through brute force enumeration over a wide range of risk-cutoff values.

1.1.3 Contribution

We have shown that a one-size-fits-all risk-cutoff value, such as 1/270, may not always be the best choice, depending on the preferences of women. Preference-sensitive risk-cutoff values for DS screening have the potential to improve the pregnancy outcomes and patient satisfaction.

Our method of determining a threshold based on women's preference is applicable to integrated screening for DS as well as all other DS screening modalities or even screening test in other health systems as long as they utilize a screen threshold to designate a positive or negative test.

We have published a peer-reviewed paper [93] on this topic. The paper is a joint work with Dr. Aaron B. Caughey in Department of Obstetrics and Gynecology at Oregon Health & Science University. In addition, we acknowledge the suggestions

and feedbacks from Dr. Scott Grosse.

1.2 Age-based Differences in the Predictive Accuracy of a One-Size-Fits-All Risk-Cutoff Value in Prenatal Integrated Screening for Down Syndrome

This and next chapters consider the population's perspective. A one-size-fits-all cutoff value is typically used in Down syndrome screening, impacting diagnostic decisions. Prior studies reported variations in age-related detection and false positive rates for combined screening but not for integrated screening.

This chapter is to assess variation in detection and false positive rates and adverse pregnancy outcomes across different age groups when a one-size-fits-all risk-cutoff value, such as 1/270, is used in integrated screening for Down syndrome. A Monte Carlo simulation was utilized to estimate the detection and false positive rates as well as adverse pregnancy outcomes. Using a one-size-fits-all risk-cutoff value, such as 1/270, can result in considerably high variations in detection and false positive rates across maternal ages, and lead to a higher than the minimum possible total number of adverse outcomes. Our findings indicate that the one-size-fits-all risk-cutoff value of 1/270, commonly used in DS screening, should be revisited and alternative (possibly age-based) cutoff values and strategies should be considered. This chapter gives rise to the question of how to select optimal risk-cutoff values, which is the focus of the next chapter.

1.3 Setting Age-specific Risk-cutoff Values in Prenatal Integrated Screening for Down Syndrome Considering the Variation of False Positive Rates across Age Groups

The use of a single risk-cutoff value in prenatal integrated screening for Down syndrome causes high variation in detection and false positive rates across maternal ages.

1.3.1 Motivation

The objective of this study is to determine the optimal age-specific risk-cutoff values that maximize the overall detection rate, while reducing the variation with respect to false positive rates across different ages.

1.3.2 Methodology

We present a flexible mathematical modeling-based framework to determine the optimal age-specific and age-group-specific risk-cutoff values. While the age-specific policy is more flexible and can result in better outcomes, the age-group-specific policy may be easier to implement. We utilize Monte Carlo simulations to assess the detection and false positive rates (and corresponding adverse pregnancy outcomes) under different risk-cutoff value choices. We present examples of policies for age-specific and age-group-specific risk-cutoff values, each resulting in fewer total number of adverse outcomes and a lower false positive rate in women of ages 36-50, compared with the commonly used single risk-cutoff value of 1/270.

1.3.3 Contribution

Age-specific risk-cutoff values have the potential to simultaneously improve pregnancy outcomes and variation false positive rates across age groups in prenatal integrated screening for Down syndrome.

1.4 Capacity Allocation in a School-Based Asthma Care Model for Pediatric Patients

Asthma is the most common childhood chronic illness [23] which affects 6.3 million (8.6% of) American children [64]. In addition, asthma is among the leading causes of severe health outcomes for children, including hospitalizations, emergency department visits, and missed school days [60]. During 2013, asthma resulted in 2,742 hospitalizations and 26,302 emergency room visits among children in Georgia [44]. About 1 in

2 children with asthma limited their usual activities, and 58% of school-aged asthma patients missed at least one school day due to asthma [44]. In fact, asthma can be well controlled by appropriate self-managements. The large amount of severe health outcomes nationwide and in Georgia highlights the lack of access and adherence to disease management care for childhood asthma.

Children's Healthcare of Atlanta (Children's) at Hughes Spalding is planning a school-based program with one mobile asthma clinic to improve the access and adherence to asthma care by outreaching underserved school-aged asthma patients in metro Atlanta. The 40-foot long, 8-foot wide vehicle (mobile asthma clinic) will be staffed by one physician, two licensed practical or registered nurses, and one drive. All the services they provide, such as diagnosis and medication, are free of charge to patients in this program. Similar programs have been established for more than a decade in several metropolitan areas in the United States and have shown improvements in patients' health outcomes.

1.4.1 Objectives

The service capacity for a mobile asthma clinic is limited to 16 patients a day and 4 days a week. Therefore, it is critical to determine how to allocate capacity among and within schools in metro Atlanta. The objective of this study is to maximize the aggregate quality of life (QoLs) scores of the asthma patient population by making two-level decisions: (i) which schools to visit, and (ii) which groups of patients within the schools to schedule for a limited number of appointments.

1.4.2 Methodology

We focus on building a mathematical framework to determine the capacity allocation. In particular, we propose a finite-horizon dynamic programming model and formulate a mixed-integer programming (MIP) for solving this model. In addition, we develop two heuristic algorithms, which can be easily implemented in practice. We

parameterize our model using the data of a local public school district and extensive computational experiments show that the two heuristics are computationally-efficient and competitive in the quality of solutions.

1.4.3 Contribution

First, to our best knowledge, we are the first in literature to consider the two-level operational decisions integrated with clinical models of disease progression, and tactically address the capacity allocation problem. The solutions we proposed can not only establish monthly appointment schedules within a certain school but also determine the capacity allocation among different schools by considering the disease dynamics in each school. Second, we formulate a MIP by considering operational constraints and disease progressions for solving the finite-horizon dynamic programming model. Using efficient optimization solvers, we can quantify the decisions for school systems of a fairly large size. Third, we develop two heuristic solutions. One is an index-based greedy algorithm, which is computationally efficient and easily implementable. The other heuristic algorithm is to solve a series of nested versions of MIP relaxation models, which requires more memory and computational time but results in a negligible gap from the optimal solutions of the original MIP. Fourth, we parameterize our model using the data of a local public school district of 89 schools and 5975 asthma patients. Our extensive computational experiments suggest that although it is intuitive to prioritize treatment of patients who are in the worst illness states, it is in fact more beneficial to appropriately divide capacity between treatment and prevention. In other words, while treating the sickest patients, we also want to prevent those patients in slightly better illness states from deteriorating to worse illness states. The findings can be applied in practice to determine the optimal schedule of the school-based asthma program.

CHAPTER II

PREFERENCE-SENSITIVE RISK-CUTOFF VALUES FOR PRENATAL-INTEGRATED SCREENING TEST FOR DOWN SYNDROME

2.1 Introduction

Down syndrome (DS, Trisomy 21) is the most common chromosomal abnormality among live births and one of the leading causes of mental retardation. The prevalence of DS in the United States increased from 9.5 to 14 per 10,000 live births between 1979 and 2006 [25, 26], despite an increasing array of screening tests during that time. Each year, about 1 of every 691 babies is born with DS [71]. Population prevalence of DS in the United States is about 8 per 10,000 people, according to a study [73] published in 2013.

For a pregnant woman considering prenatal screening for DS, there are at least two major adverse outcomes of interest: (i) *undetected DS live birth*, that is, birth of a baby with undetected DS prior to delivery; and (ii) *euploid procedure-related fetal loss*, due to an invasive diagnostic test, such as amniocentesis. The likelihoods of these two outcomes depend on the risk-cutoff value used in recommending amniocentesis. A low risk-cutoff value results in a high detection rate but also leads to a higher false-positive rate, giving rise to more unnecessary amniocenteses and euploid procedure-related fetal losses. On the other hand, a higher risk-cutoff value results in fewer avoidable fetal losses but leads to a higher false-negative rate and more undetected DS live births. Currently, about one to five out of every 1000 pregnant women undergoing amniocentesis experience a procedure-related fetal loss [18, 67].

Evidence suggests that women have different preferences about the pregnancy outcomes. For example, cross-sectional studies of diverse pregnant women showed that substantial differences exist in preferences with respect to the two adverse outcomes [55, 57, 58]. Another study showed that women who perceive themselves at higher risk or are more concerned about having a DS live birth may prefer to maximize detection rate at the expense of a higher false-positive rate, while those who are more concerned about a procedure-related fetal loss may prefer to minimize the risk of an avoidable fetal loss at the expense of a higher false-negative rate [92]. Despite the existing evidence that women have different preferences, no preference-based guideline of DS screening is in place [56, 59].

There are several studies [11, 17, 19, 20, 49, 68] utilizing quality-adjusted life years to capture womens preferences in cost-effectiveness analysis. However, these studies commonly use 1/270 as the risk-cutoff value.[11] Walker et al.[92] recently acknowledged that different women may weigh pregnancy outcomes differently and considered the effects of such differences in choosing the risk-cutoff values for quadruple serum test (quad screen). Compared with the quad screen, *integrated screening* has a higher detection rate.[1, 84] However, the question of how the risk-cutoff values should be set in *integrated screening* to capture womens various *risk preferences* remains unanswered.

The purpose of this study is to *determine the optimal risk-cutoff value* of integrated screening by considering *risk preferences of women* with respect to pregnancy outcomes. Based on a mathematical model, we (i) estimated the likelihoods of undetected DS live births and euploid procedure-related fetal losses over a wide range of risk-cutoff values; (ii) explored the conditions under which a certain risk-cutoff value became optimal; and (iii) quantified the impacts of using the optimal risk-cutoff values on pregnancy outcomes with respect to womens preferences.

2.2 Methods

2.2.1 Screening strategy

Our study focused on the integrated screening strategy, which assesses the DS risk by synthesizing biomarkers in the first trimester (nuchal translucency, pregnancy associated plasma protein A, and free β human chorionic gonadotropin) and the second trimester (alpha-fetoprotein, human chorionic gonadotropin, unconjugated estriol, and Inhibin-A) with maternal age.[89] Those women who have a higher calculated risk of DS compared with a pre-defined risk-cutoff value of integrated strategy are offered the option of diagnostic tests, which might lead to procedure-related fetal losses. We assumed that women with a lower risk of DS do not consider any diagnostic tests. Note that DS pregnancies might also result in spontaneous miscarriages or stillbirths at a higher rate than euploid pregnancies.[61] Figure 1 depicts the integrated screening strategy and the resulting pregnancy outcomes.

2.2.2 Decision model

We built a mathematical model to quantify the optimal risk-cutoff values given the womens preferences, represented by different weights (penalties) assigned to different pregnancy outcomes: W_1 for an undetected DS live birth and W_2 for a euploid procedure-related fetal loss. We focused on the relative weight, that is, $W_1 : W_2$, rather than absolute values. For example, if an undetected DS live birth is twice as undesirable for a woman as a euploid procedure-related fetal loss, then we have $W_1 : W_2 = 2 : 1$. We tested various values for $W_1 : W_2$ ratio to quantify the impact of different relative preferences on the optimal risk-cutoff values.

We used Monte Carlo simulation [11, 17, 89, 90] to assess the probability of an undetected DS live birth, probability of a procedure-related fetal loss, detection rate (DR), and false-positive rate (FPR) for various cutoff values ranging from 1/2000 to 1/2. Because the risks of DS are different across age groups, we reported performance

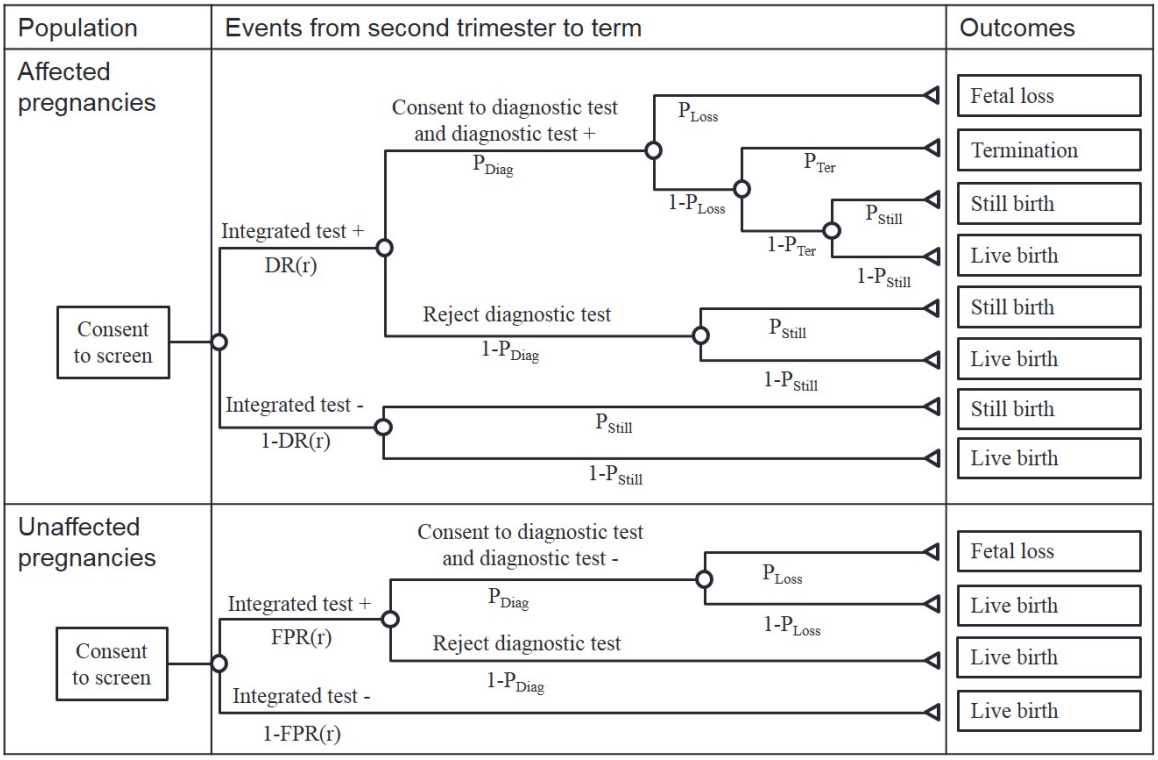

Figure 1: Algorithm for integrated strategy from the second trimester to term. The second-trimester diagnostic test is amniocentesis. r, the risk-cutoff value used; DR(r), detection rate for the risk-cutoff value r; FPR(r), false-positive rate for the risk-cutoff value r; P_{Diag}, consent rate to diagnostic test given positive results of screening tests; P_{Loss}, procedure-related fetal-loss rate due to amniocentesis; P_{Ter}, termination rate given positive results of diagnostic tests; P_{Still}, spontaneous fetal-loss rate of DS pregnancies.

measures for the overall population and selected age groups (20, 30, 35, and 40 years old), and compared different risk-cutoff values within certain age groups.

2.2.3 Data and model parameters

Based on the maternal age distribution from the United States National Vital Statistics birth data between years 2008 and 2010,[29] we sampled 100,000 affected pregnancies and unaffected pregnancies, respectively, to calculate DR and FPR of integrated screening test for each age group. For a given risk-cutoff value, to estimate DR, we divide the number of affected cases with a risk exceeding the risk-cutoff value by 100,000. Similarly, to estimate FPR, we divide the number of non-affected cases with a risk exceeding the risk-cutoff value by 100,000. Marker levels (in multiples of median among unaffected pregnancies) of integrated screening tests were sampled based on the distributions reported in the Serum, Urine, and Ultrasound Screening Study (SURUSS)[90, 91] using MATLAB (The MathWorks Inc., Natick, MA, 2012). We calculated the risk of DS using the standard method of Wald and Hackshaw (2000).[88] In addition, we assumed 100% sensitivity and specificity of amniocentesis.[34, 51] Table 1 presents the model parameters and corresponding sources.

2.2.4 Sensitivity analysis

We conducted one-way sensitivity analysis on a consent rate to amniocentesis given positive results of screening tests (P_{Diag}) and procedure-related fetal loss rate due to amniocentesis (P_{Loss}), and we solved for the optimal risk-cutoff values for the extreme values (i.e. lower and higher levels) of these two parameters. In addition, we also conducted probabilistic sensitivity analysis for the two rates based on 1,000 samples drawn from triangular distributions representing the uncertainty in each rate. The lower limit, upper limit, and mode of the triangular distribution for P_{Diag} were 50%, 70%, and 57.1%, respectively.[11] Similarly, the lower limit, upper limit, and mode of the triangular distribution for P_{Loss} are 0.1%, 0.5%, and 0.3%, respectively.[18, 67]

Table 1: Values of parameters.

Parameters	Explanation	Value for baseline scenario (range)	References
P_{Diag}	Consent rate to amniocentesis given positive results of screening tests	$57.1\%(50-70\%)$	Ball et al. (2007) [11]
P_{Ter}	Termination rate given positive results of diagnostic tests	60%	Estimated based on Natoli et al. (2012) [66]
P_{Still}	Spontaneous fetal loss rate due to Down syndrome	23%	Morris et al. (1999)[61]
P_{Loss}	Procedure-related fetal loss rate due to amniocentesis	$0.3\%(0.1-0.5\%)$	Caughey et al. (2006) [18]; Odibo et al. (2008) [67]
$W_1 : W_2$	Ratio of penalties for a Down Syndrome live birth versus a euploid procedure-related fetal loss	2:1, 1:1, 1:2	We vary this ratio to test the impact of preferences on optimal risk-cutoff values

2.3 Results

Table 2 and Figure 2 summarize the performance measures for various risk-cutoff values, including DR, FPR, probabilities of undetected DS live births (P_1), and euploid procedure-related fetal losses (P_2). In the overall simulated population at the second trimester, 0.237% has a DS affected fetus, which is in line with the DS prevalence of 1 in 420 in Walker et al.[92] As the risk-cutoff value decreases from 1/50 to 1/1000, DR increases from 82.0% to 95.8% and FPR increases from 0.7% to 8.8%. For a particular risk-cutoff value (e.g. 1/270), the DR and FPR are lower in younger age groups (e.g. 83.5% and 1.5%, respectively, for 20-year-old women) than those in older age groups (e.g. 96.6% and 15.4%, respectively, for 40-year-old women). As expected, probabilities of an undetected DS live birth and a procedure-related fetal loss are also lower in younger age groups.

The optimal risk-cutoff values for the overall population and each age group are presented in Table 3. When the relative penalty a woman places on an undetected DS live birth (W_1) versus a euploid procedure-related fetal loss (W_2) is the following: (i) higher ($W_1 : W_2$=2:1), the optimal risk-cutoff value is 'low', 1/915; (ii) the same ($W_1 : W_2$=1:1), the optimal risk-cutoff value is 'medium', 1/454; (iii) lower ($W_1 : W_2$=1:2), the optimal risk-cutoff value is 'high', 1/237. For the overall population or a given age group, as the relative preferences change, the optimal risk-cutoff value changes significantly. For example, for age 30 years, for the relative penalties ($W_1 : W_2$) of 2:1, 1:1, and 1:2, the optimal risk-cutoff values are 1/864, 1/451, and 1/238, respectively. Furthermore, for a specific combination of age group and preferences, the optimal risk-cutoff value can range from 1/994 to 1/221.

Depending on the optimal risk-cutoff values reflecting womens preferences, the proportion of the total number of adverse outcomes may change. For example, if the 1/270 threshold was applied to a cohort of 4 million women (approximately the annual number of live births in the United States [29]), there would be about 636

Table 2: Measured outputs at various pre-defined risk-cutoff values. P_1, likelihood of an undetected Down Syndrome live birth in a cohort of pregnant women (entire population or a certain age group); P_2, likelihood of a procedure-related fetal loss in a cohort of pregnant women (entire population or a certain age group); DR, detection rate of integrated screening test, the proportion of affected pregnancies with positive results of screening tests; FPR, false-positive rate of integrated screening test, the proportion of unaffected pregnancies with positive results of screening tests.

Risk-cutoff value	Age	P_1(%)	P_2(%)	DR(%)	FPR(%)
1/270	Overall	0.016	0.005	91.3	3.2
	20	0.011	0.003	83.5	1.5
	30	0.014	0.004	86.4	2.3
	35	0.024	0.009	91.6	5.4
	40	0.040	0.026	96.6	15.4
1/50	Overall	0.033	0.001	82.0	0.7
	20	0.020	0.001	70.4	0.3
	30	0.027	0.001	74.2	0.4
	35	0.052	0.002	81.7	1.2
	40	0.113	0.007	90.2	4.3
1/500	Overall	0.011	0.009	93.7	5.3
	20	0.009	0.005	87.3	2.7
	30	0.011	0.007	89.8	4.0
	35	0.017	0.015	94.1	8.8
	40	0.024	0.038	97.9	22.8
1/1000	Overall	0.008	0.015	95.8	8.8
	20	0.006	0.008	91.0	5.0
	30	0.007	0.012	93.0	7.1
	35	0.010	0.024	96.3	14.3
	40	0.013	0.056	98.9	33.2

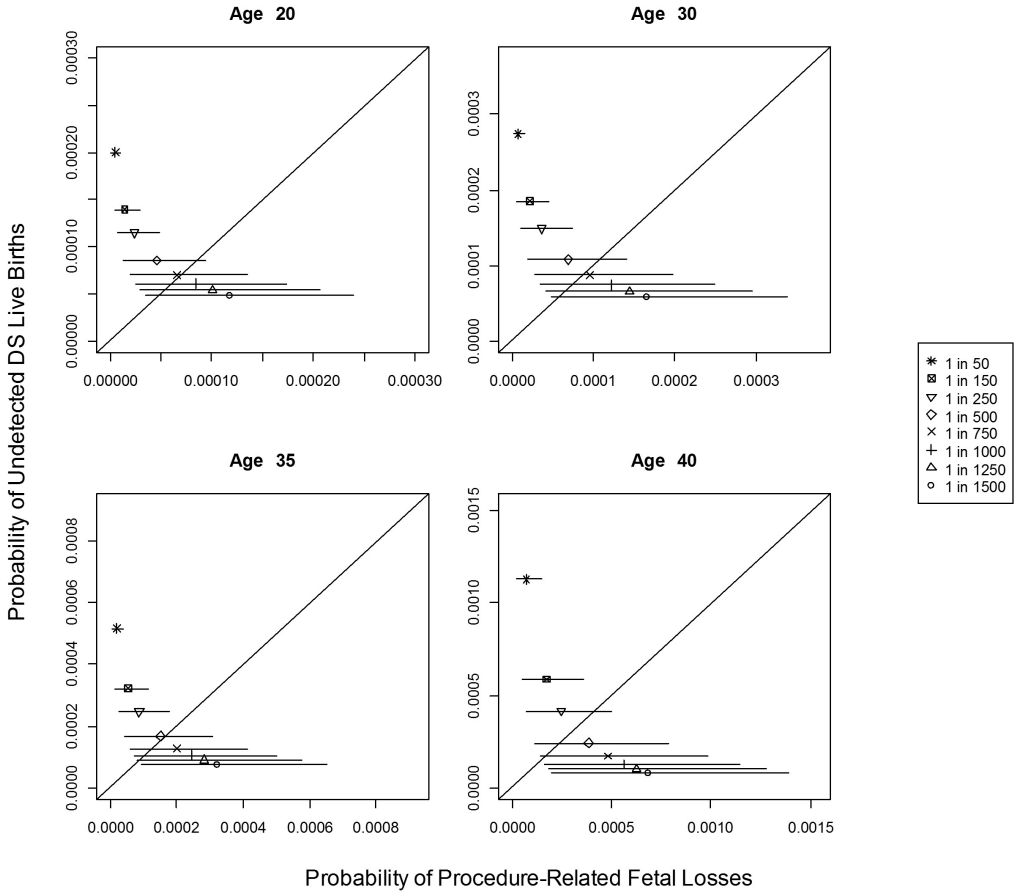

Figure 2: Probabilities of euploid procedure-related fetal losses and undetected DS live births for various age groups as a function of the risk-cutoff value. Each horizontal line segment at each point denotes the possible range of the probability of procedure-related fetal losses when P_{Loss} (0.1−0.5%) and P_{Diag} (50−70%). Probability of procedure-related fetal losses: proportion of procedure-related fetal losses in a maternal age group (i.e. the number of procedure-related fetal losses divided by the total number of pregnancies in that maternal age group); Probability of undetected DS live births: proportion of undetected DS live births in an age group (i.e. the number of undetected DS live births divided by the total number of pregnancies in that maternal age group).

Table 3: Optimal risk-cutoff values for various preferences and age groups. P_1, likelihood of an undetected Down Syndrome live birth in a cohort of pregnant women (entire population or a certain age group); P_2, likelihood of a procedure-related fetal loss in a cohort of pregnant women (entire population or a certain age group); DR, detection rate of integrated screening test, the proportion of affected pregnancies with positive results of screening tests; FPR, false-positive rate of integrated screening test, the proportion of unaffected pregnancies with positive results of screening tests; Proportion of women with amniocentesis: number of women who undergo amniocentesis divided by total number of women in a cohort; Proportion of women with positive amniocentesis results: number of women who undergo amniocentesis and receive positive diagnostic results divided by total number of women in a cohort.

$W_1 : W_2$	Age	Optimal risk-cutoff value	P_1 (%)	P_2 (%)	DR (%)	FPR (%)	Proportion of women with amniocentesis (%)	Proportion of women with positive amniocentesis results (%)
Higher 2:1	Overall	1/915	0.008	0.014	95.6	8.2	4.83	0.13
	20	1/898	0.006	0.008	90.5	4.5	2.62	0.05
	30	1/864	0.008	0.011	92.4	6.3	3.66	0.07
	35	1/912	0.011	0.023	96.1	13.4	7.83	0.20
	40	1/994	0.013	0.056	98.9	33.1	19.45	0.85
The same 1:1	Overall	1/454	0.012	0.008	93.4	4.9	2.90	0.13
	20	1/430	0.009	0.004	86.5	2.3	1.36	0.04
	30	1/451	0.011	0.006	89.3	3.7	2.16	0.07
	35	1/472	0.017	0.014	93.9	8.5	5.01	0.20
	40	1/446	0.027	0.036	97.7	21.3	12.80	0.84
Lower 1:2	Overall	1/237	0.017	0.005	90.7	2.9	1.76	0.12
	20	1/235	0.012	0.002	82.6	1.3	0.79	0.04
	30	1/238	0.015	0.003	85.7	2.0	1.22	0.07
	35	1/231	0.026	0.008	90.9	4.8	2.92	0.19
	40	1/221	0.045	0.023	96.1	13.4	8.38	0.83

undetected DS live births (resulting from 825 false-negative screening tests) and 219 euploid procedure-related fetal losses, that is, a total of 855 adverse outcomes. If a risk-cutoff value of 1/454 was used instead, there would be about 485 undetected DS live births and 333 euploid procedure-related fetal losses, a total of 818 adverse outcomes. Although 1/454 increases the number of euploid procedure-related fetal losses by 114, it decreases the number of undetected DS live births by 151 and the total number of adverse outcomes is reduced by 37. For those women who treat the two outcomes as equally undesirable, the 1/454 risk cutoff would be preferable to the 1/270 cutoff value, which is commonly used in practice.

2.3.1 Sensitivity analysis

Table 4 presents the results of a one-way sensitivity analysis of two parameters: consent rate to amniocentesis given positive results of screening tests (P_{Diag}) and procedure-related fetal loss rate due to amniocentesis (P_{Loss}). Higher P_{Diag} and P_{Loss} result in higher optimal risk-cutoff values. The results are sensitive to P_{Diag} and P_{Loss}. When women weigh the two outcomes equally ($W_1 : W_2$=1:1), the optimal risk-cutoff value can vary from 1/1561 to 1/225 depending on P_{Diag} and P_{Loss} (Table 4). Figure 3 shows the probabilistic sensitivity analysis results for P_{Diag} and P_{Loss}. Similar to one-way sensitivity analysis, these results confirm that the optimal risk-cutoff values are sensitive to these two parameters.

Table 4: Sensitivity analysis of optimal risk-cutoff values at different values of P_{Diag} and P_{Loss} when $W_1 : W_2$=1:1.

		P_{Diag}		
Optimal risk-cutoff value		50%	57.1%	70%
P_{Loss}	0.5%	1/330	1/279	1/225
	0.3%	1/532	Base case: 1/454	1/374
	0.1%	1/1561	1/1336	1/1113

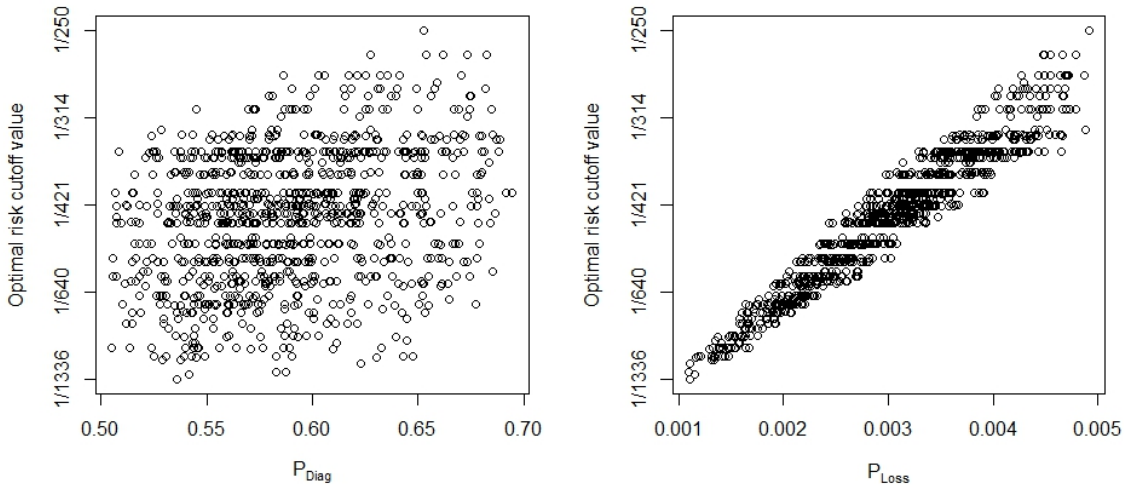

Figure 3: Probabilistic sensitivity analysis of P_{Diag} and P_{Loss} based on 1000 samples when $W_1 : W_2$=1:1. P_{Diag}, consent rate to diagnostic test given positive results of screening tests; P_{Loss}, procedure-related fetal loss rate due to amniocentesis.

2.4 Discussion

We proposed a comprehensive model to select risk-cutoff values for the integrated screening test.

Our analyses showed that if women have varying preferences, a one-size-fits-all risk-cutoff value, such as 1/270, is not always the best choice, depending on the preferences of women. For example, for some women, if an undetected DS live birth is significantly more undesirable than a euploid procedure-related fetal loss,[55, 57, 58] then a lower cutoff value would be more appropriate. Our proposed method ensures that the cutoff value balances the risks of the adverse pregnancy outcomes depending on a group of womens preferences; this is a step forward to personalized decisions.

The overall detection and false-positive rates cannot be improved simultaneously by applying age-specific risk-cutoff values compared with using one risk-cutoff value for all age groups. This may be due to the similarity of the optimal risk-cutoff values

across different age groups, if preferences of pregnancy outcomes are the same for all women (shown in Table 3).

The sensitivity analysis indicates that the optimal risk-cutoff value could vary significantly for different values of positive results of screening tests (P_{Diag}) and the procedure-related fetal loss rate due to amniocentesis (P_{Loss}). In particular, we find that as P_{Diag} or P_{Loss} increases, using higher cutoff values would be more appropriate. While we used population level average for P_{Diag}, we acknowledge that it actually varies significantly from one woman to another. Especially for those women who are more likely to consent to amniocentesis, a higher cutoff value may reduce the overall rate of procedure-related fetal losses.

Despite the novel results of our study, it is not without limitations. Our results are based on mathematical models with assumptions, not randomized controlled trials (RCTs) or observational studies. For example, we treated the DS spontaneous miscarriage rate (P_{Still}) independent of the result of the integrated test, while the SURUSS study [90, 91] noted that those pregnancies that end in spontaneous miscarriages tend to receive a higher risk score in screening tests. We also assumed that women with a predicted risk of DS below the risk-cutoff value would not receive a diagnostic test. Although imperfect, our mathematical model allows for testing a spectrum of risk-cutoff values and differences in womens preferences, which is not possible in RCTs and observational studies. Another limitation is using the distributions of marker levels from the SURUSS study in the UK, similar to other studies which used the SURUSS trial for US-based [92] and Canada-based [39, 40] analysis. Hence, the parameters (including means, standard deviations, correlation coefficients, and truncation limits) may not well present the characteristics of women in the United States or other countries. Decision-makers in the Unites States or other countries can adjust marker levels based on local data to choose appropriate risk-cutoff values.

In this study, we only focus on the integrated screening test and do not consider the cell-free fetal DNA (cffDNA) test, a new method for DS screening that has a higher detection rate and a lower false-positive rate than traditional screening methods.31,32 [2, 35] Currently, cffDNA is not recommended for low-risk women.[2] In current practice, ultrasound/serum screenings, including integrated, sequential, contingent, quadruple, or combined tests, are still the most commonly used strategies. However, our modeling framework is general and can easily be extended to model other screening strategies and capture the improvements in detection and false-positive rates. Further, given that any of these screening modalities utilize a screen threshold to designate a positive or negative test, the findings that such a threshold should be based upon women's preferences are applicable.

2.5 Conclusion

Prenatal screening decisions need to take into account patient preferences for different pregnancy outcomes when determining the risk-cutoff value. The model in our paper can be used to identify appropriate risk-cutoff values of integrated screening based on women's preferences.

CHAPTER III

AGE-BASED DIFFERENCES IN THE PREDICTIVE ACCURACY OF A ONE-SIZE-FITS-ALL RISK-CUTOFF VALUE IN PRENATAL INTEGRATED SCREENING FOR DOWN SYNDROME

3.1 Introduction

Down syndrome (DS) is among the most prevalent genetic disorders in human beings. Maternal serum and ultrasound screening are standard noninvasive prenatal screening strategies administered to pregnant women to assess the risk of having a baby with DS. Acknowledging the clinical availability and high accuracy of cell-free DNA (cfDNA) screening, the American College of Obstetricians and Gynecologists and the Society for Maternal-Fetal Medicine still suggest conventional screening methods [3] as the most appropriate choice for most pregnant women. Among conventional screening options for DS, this study focuses on prenatal integrated screening, which combines both first and second trimester biomarkers [11] and has the highest detection rate (DR) [1], compared to the first trimester combined screening and the second trimester triple and quadruple screenings.

There are two major adverse outcomes of interest in DS screening: (1) an undetected DS case prior to delivery, which is referred to as *undetected DS live birth*, and (2) a *euploid procedure-related fetal loss* due to an invasive diagnostic test such as amniocentesis, following a DS screening test with "positive" result. The likelihoods of these two outcomes depend on the risk-cutoff value used in the screening test, and the corresponding test accuracy. In practice, a one-size-fits-all type of risk-cutoff value,

such as 1/270 [11], is typically used to designate the outcomes of a maternal serum screening, which measures the DS risk as a continuous variable, as positive and negative. "Detection rate" (or equivalently sensitivity) and "false positive rate" (FPR, or equivalently 1specificity) refer to the percentage of pregnancies labeled as positive by the test, correctly and incorrectly, respectively. "False negative rate" (FNR) is defined as 1−DR. Hence, a lower DR (equivalently, higher FNR) increases adverse outcome (1), that is, undetected DS live births, and a higher FPR increases adverse outcome (2), that is, euploid procedure-related fetal losses.

Several studies have reported differences in preferences among women regarding the two main adverse pregnancy outcomes [55, 57, 58, 92, 93]. In addition, the risk of DS [62] changes significantly with age. This raises the question of whether the commonly used risk-cutoff value of 1/270 is indeed the best option for all ages.

The purpose of this study is to assess the variation in detection and false positive rates (and hence, the undetected DS live births and euploid procedure-related fetal losses) across different age groups when a one-size-fits-all risk-cutoff value, such as 1/270, is used in integrated screening for DS. Prior researches [38, 69, 83, 85, 93] acknowledged variations in age-related detection and false positive rates in the context of first and/or second trimester testing. Among them, Yan et al. [93] is particularly relevant, which, by taking individual patients perspective, showed that the optimal risk-cutoff values could be very different depending on womens personal preferences.

However, to our knowledge, no prior population-based study has considered the effect of using a single cutoff value in integrated screening on the health outcomes disparities with respect to detection and false positive rates across different age groups, which is the focus of this study.

3.2 Methods

Prenatal integrated screening test considers six biomarkers [11, 89] including nuchal translucency, pregnancy associated plasma protein A (PAPP-A) in the first trimester, and alpha-fetoprotain, human chorionic gonadotropin (hCG), unconjugated estriol, and Inhibin-A in the second trimester. We used a previously validated Monte Carlo simulation [93] to sample from the biomarker distributions reported in the Serum, Urine, and Ultrasound Screening Study (SURUSS) [91]. The risk of a DS pregnancy is calculated using the biomarkers and maternal age [88]. In particular, we sampled 800,000 affected pregnancies, calculated the risk for each pregnant woman, counted the number of pregnant women with a risk higher than the given risk-cutoff value (i.e., the pregnancies identified as positive by the test), and divided this number by 800,000 to derive DR. Similarly, we sampled 800,000 unaffected pregnancies, calculated the risk for each pregnant woman, counted the number of pregnant women with a risk higher than the given risk-cutoff value, and divided this number by 800,000 to derive FPR. Similar to Yan et al. [93], we used a Monte Carlo simulation to calculate DRs and FPRs for risk-cutoff values: 1/63, 1/137, 1/270, 1/397, 1/485, 1/536, 1/1052, across different ages. 1/270 is a commonly used risk-cutoff value in practice and all the values listed here result in FPRs between 1% and 10%. In addition, we also reported the total number and probabilities of adverse outcomes, i.e., undetected DS live births (DSLs), and euploid procedure-related fetal losses (EFLs) due to amniocentesis, respectively, out of a cohort of 4 million pregnancies (approximately the annual size of live births in the United States [87]). The parameters [93] to calculate the adverse outcomes include uptake rate of amniocentesis (57.1%), spontaneous fetal loss rate due to DS (23%), and procedure-related fetal loss rate due to amniocentesis (0.3%).

3.3 Results

Table 5 and Figure 4 summarize the overall DRs and FPRs of various one-size-fits-all risk-cutoff values. Figure 5 illustrates the tradeoff between probabilities of DSLs and EFLs. The expected numbers of DS live births and euploid procedure-related fetal losses are presented in Table 5 and Figure 6. Using the risk-cutoff value of 1/270 results in an overall DR of 91.01%, an overall FPR of 3.56%, and a total number of 937 adverse outcomes. For cutoff value 1/270, the DR and FPR for each maternal age are presented in Figure 7. The FPRs range from 1.59% to 33.8% and DRs range from 82.60% to 98.73% for ages 20–50. In addition, we present age-based FPRs and DRs for various alternative risk-cutoff values in Figures 8–9.

Table 5: Overall DRs and FPRs, DSLs, and EFLs of 1/270 and alternative one-size-fits-all risk-cutoff values. Note that 1/270 is highlighted. Among the alternative risk-cutoff values, 1/397, 1/485, and 1/536 results in the same total number of DSLs and EFLs, which is smaller than that of 1/270. However, the numbers of DSLs and EFLs are different under each of the three risk-cutoff values. DRs, detection rates; FPRs, false positive rates; DS, Down syndrome; DSLs, undetected DS live births; EFLs, euploid procedure-related fetal losses.

	One-size-fits-all risk-cutoff values						
Model outputs	1/63	1/137	1/270	1/397	1/485	1/536	1/1052
Overall DR (%)	83.03	87.65	91.01	92.64	93.40	93.76	95.86
Overall FPR (%)	1.00	2.00	3.56	4.85	5.67	6.12	10.00
DSLs	1319	956	696	568	511	483	319
EFLs	66	135	241	330	387	415	683
DSLs + EFLs	1385	1091	937	898	898	898	1002

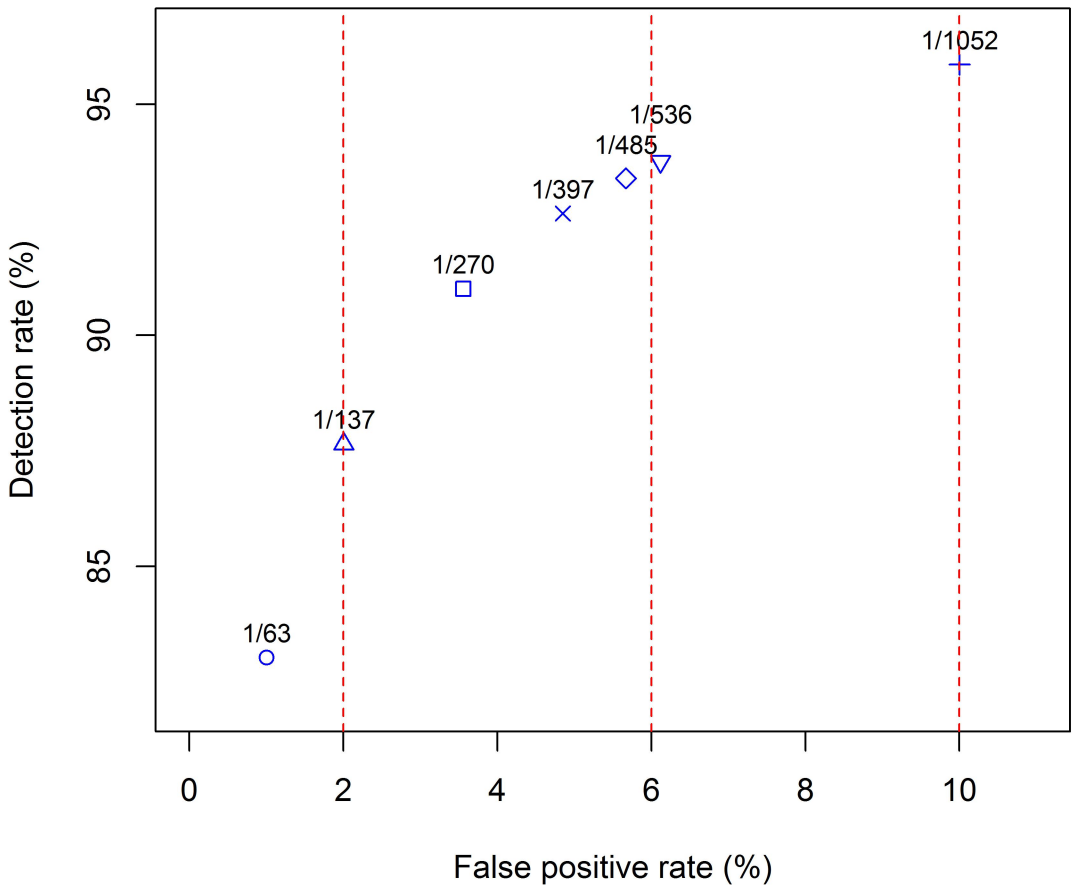

Figure 4: Overall detection and false positive rates (DRs and FPRs) of 1/270 and alternative one-size-fits-all risk-cutoff values.

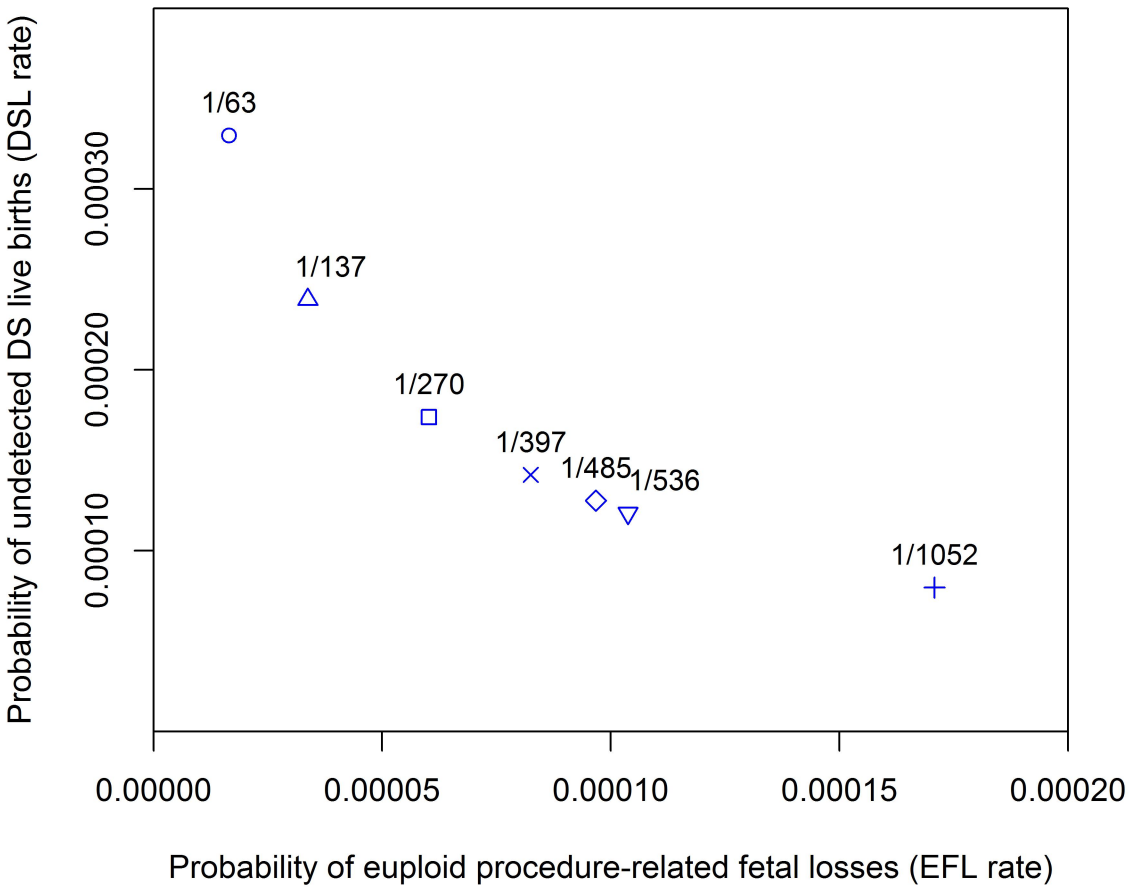

Figure 5: Probabilities of undetected Down syndrome live births (DSLs) and euploid procedure-related fetal losses (EFLs) of 1/270 and alternative one-size-fits-all risk-cutoff values.

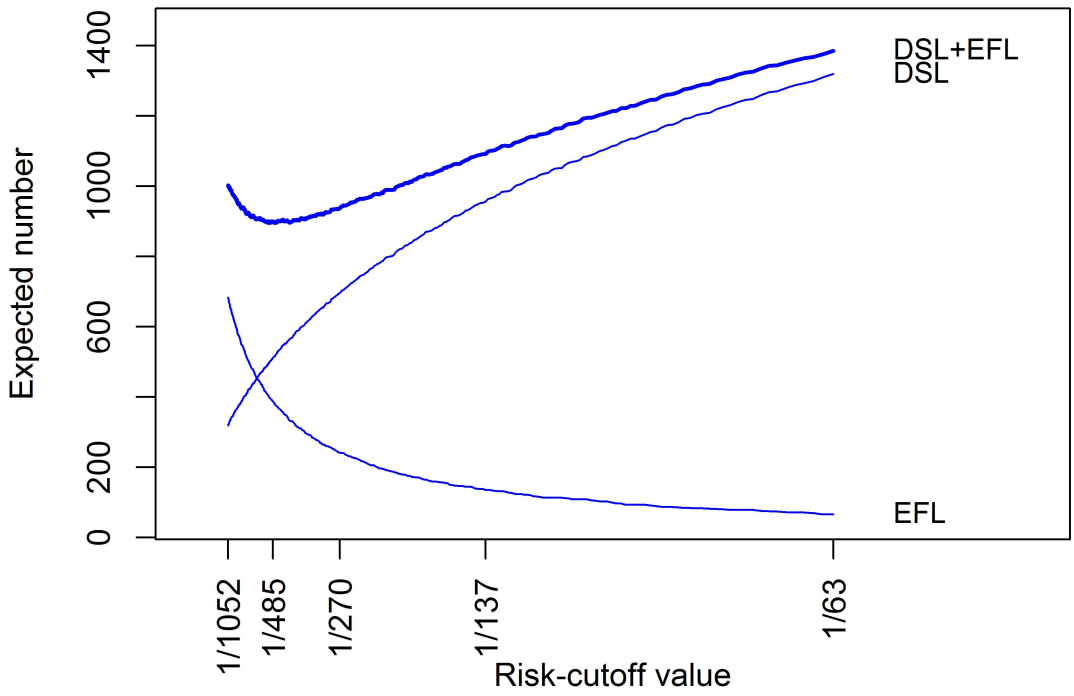

Figure 6: Number of undetected Down syndrome live births (DSLs), euploid procedure-related fetal losses (EFLs), and total adverse outcomes of one-size-fits-all risk-cutoff values between 1/1052 and 1/63.

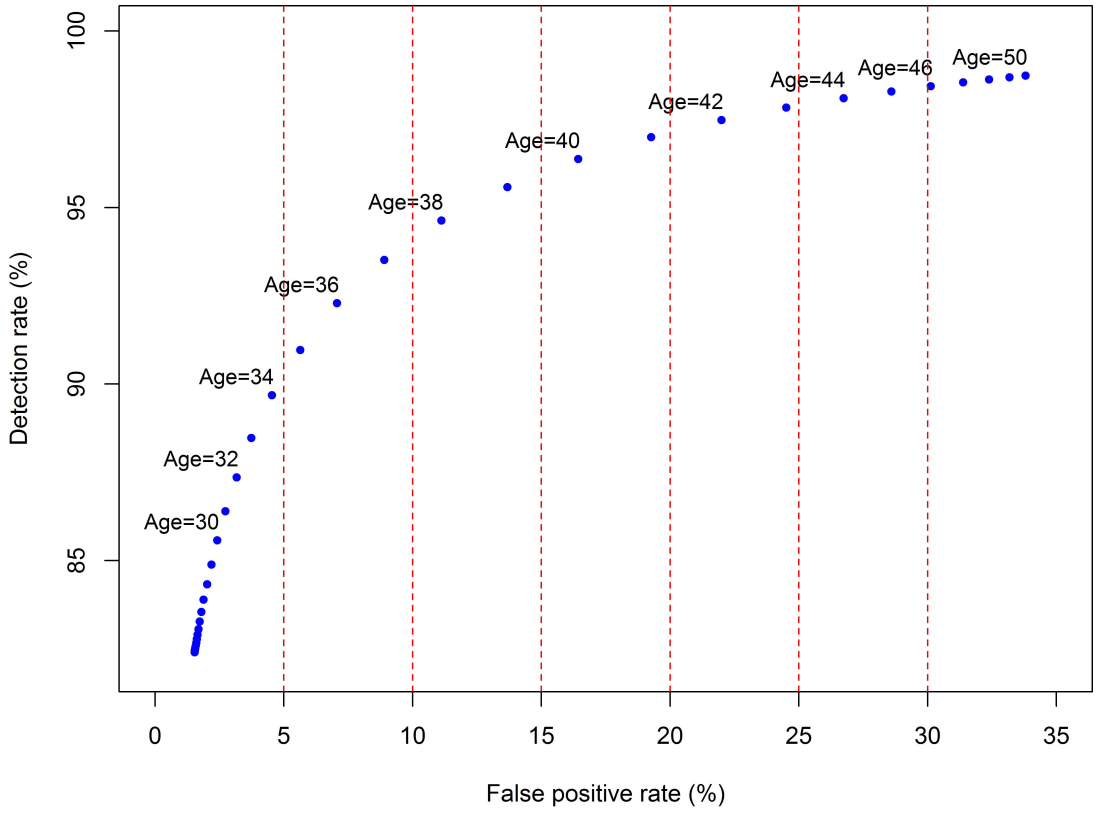

Figure 7: Detection and false positive rates across maternal ages (blue dots) using the risk-cutoff value of 1/270. The vertical (red) dashed lines show the level of false positive rates of 5%, 10%, 15%, 20%, 25%, and 30%.

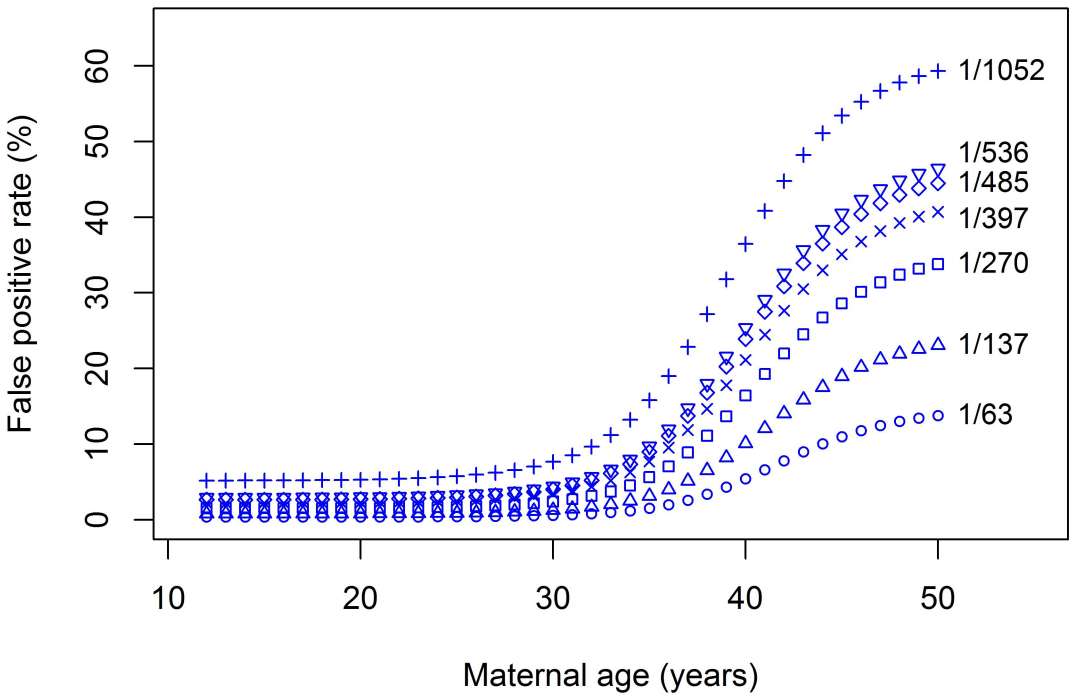

Figure 8: False positive rates in individual ages under different risk-cutoff values.

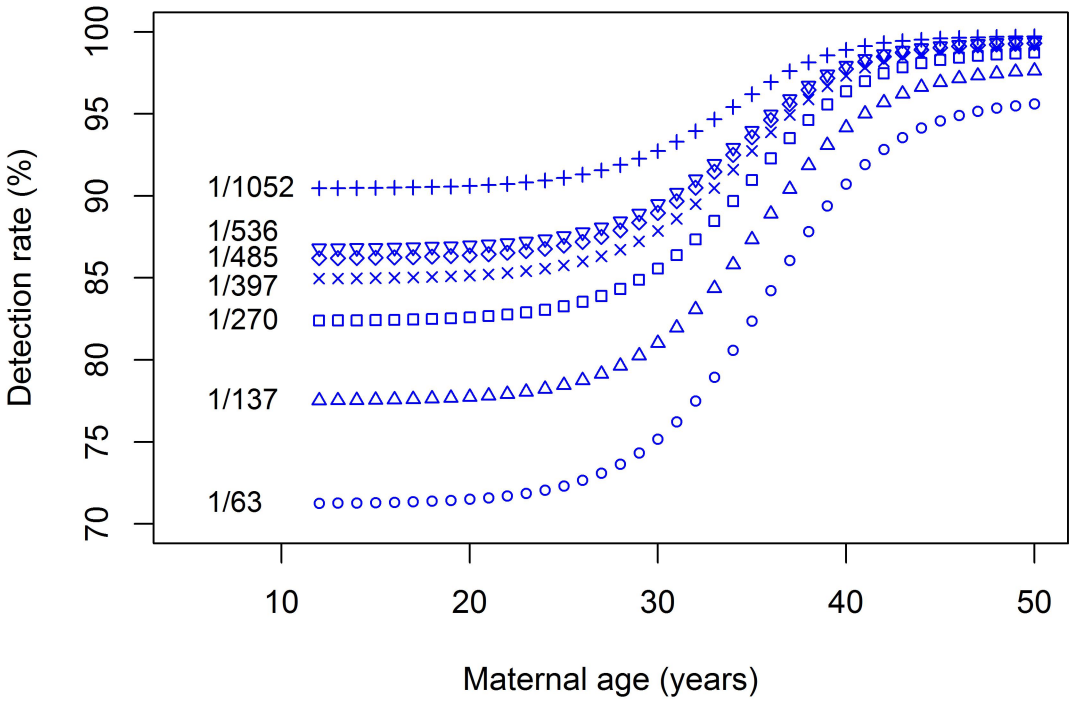

Figure 9: Detection rates in individual ages under different risk-cutoff values.

3.4 Discussion

3.4.1 Considerable variation in DR and FPR across ages

Figures 8-9 show that when a single cutoff value is used, the DRs and FPRs (and hence, adverse pregnancy outcomes) exhibit significant variations across maternal ages. For example, with a risk-cutoff value of 1/270, the FPR for women younger than 35 years is below 5% while for women older than 45 years the FPR exceeds 30%. Regardless of the cutoff value, both the DRs and FPRs increase with maternal age (Figures 8-9). However, variation exists in womens preferences with respect to the pregnancy outcomes [57, 58, 62, 92, 93]. For example, given that the chance of a future pregnancy for older women is significantly lower, it is not clear that they would prefer this combination of high DR and high FPR (resulting from the risk-cutoff value of 1/270), since a high FPR could lead to a procedure-related fetal loss

due to a follow-up diagnostic test. Hence, it is important to consider the patient age and preferences, along with other factors, in the choice of the risk-cutoff value.

3.4.2 Number of adverse outcomes

Note that the commonly used risk-cutoff value 1/270 approximately balances the false negative rate ($100\% - 92.29\% = 7.71\%$) and FPR (7.07%) for a 36-year-old woman. Recall that the false negative rate (FNR) is $1-$DR, i.e., the percentage of DS pregnancies not detected by the test. However, 1/270 leads to an imbalance between FNR and FPR in other maternal ages; for example, for 40-year-old women the FNR is 3.62% and FPR is 16.43% (Figure 7). In addition, 1/270 does not minimize the total number of adverse outcomes, resulting in 696 undetected DS live births and 241 euploid procedure-related fetal losses, with a total of 937 adverse outcomes across all ages (Table 5). In comparison, 1/485 results in 511 undetected DS live births and 387 euploid procedure-related fetal losses, with a total of 898 adverse outcomes. In general, as the cutoff value decreases, both DR and FPR increase, i.e., the number of undetected DS live births decreases and the number of euploid procedure-related fetal losses increases. While it is not possible to improve both adverse outcomes at the same time by choosing a cutoff value other than 1/270, it is possible to change (and improve) the total number of adverse outcomes, as well as their distribution across different age groups. For example, 1/397, 1/485, and 1/536 all result in 898 total number of adverse outcomes, but the number of undetected DS live births and euploid procedure-related fetal losses are different under each of the three risk-cutoff values.

3.5 Conclusion

Currently, a single one-size-fits-all risk-cutoff value is often used to label the DS risk of a pregnancy as positive or negative based on the results of integrated screening. This study shows that the use of a single risk-cutoff value can lead to high variation

of false positive rates and detection rates across maternal ages, impacting the number of adverse outcomes (undetected DS live births and euploid procedure-related fetal losses) across ages. Further, the commonly used risk cutoff value, 1/270, does not minimize the total number of adverse outcomes. These results highlight the need for revisiting the risk cutoff values used in integrated screening, and possibly adjusting them considering the age and the preferences (regarding adverse outcomes) of pregnant women.

CHAPTER IV

SETTING AGE-SPECIFIC RISK-CUTOFF VALUES IN PRENATAL INTEGRATED SCREENING FOR DOWN SYNDROME CONSIDERING THE VARIATION OF FALSE POSITIVE RATES ACROSS AGE GROUPS

4.1 Introduction

Down syndrome (DS, Trisomy 21) is a common type of chromosomal abnormality associated with an increasing risk of several problems such as mental retardation and congenital heart diseases, with a prevalence of about 1 in every 691 births in the United States [71]. Prenatal DS screening and diagnostic decisions affect about overall 4 million live births [87].

Prenatal integrated screening, a serum and/or ultrasound markers-based non-invasive test, is commonly used to assess the risk of having a baby with Down syndrome. The result is reported as positive (negative) when the derived risk is higher (lower) than a predefined risk-cutoff value. The outcome is a false positive when the screening result is positive but the baby does not have Down syndrome; conversely, the outcome is a false negative when the screening result is negative but the baby does have Down syndrome. The use of a higher risk-cutoff value reduces the false positive rate, but increases the false negative rate, and vice versa. The false negative and false positive rates relate to adverse outcomes of (i) an undetected DS case prior to delivery, which is referred to as *undetected DS live birth*, and (ii) a *euploid procedure-related fetal loss* due to an invasive diagnostic test such as amniocentesis, following a DS screening test with a positive screening result. The detection rate refers to the

fraction of pregnancy outcomes which are identified correctly by the screening result. In current practice, a one-size-fits-all type value, typically 1/270, is commonly used for prenatal-integrated screening test for DS [11].

A recent study [94] has shown that the use of a one-size-fits-all risk-cutoff value in prenatal integrated screening may result in undesirably high false positive rates, and in turn, potentially higher rates of euploid procedure-related fetal losses, among older women.

The purpose of this study is to develop a framework to identify the optimal age-specific risk-cutoff values that maximize the overall detection rate while considering variation across different maternal ages with respect to the false positive rates. We proposed a flexible mathematical model (based on integer programming), and solved it using the state-of-the-art optimization solvers. To prepare the input to the optimization framework, we developed a simulation model to assess the performance of the computed optimal risk-cutoff values versus the commonly used one-size-fits-all risk-cutoff value.

4.2 Methods

We developed an integer programming (IP) model for determining the optimal risk-cutoff values of prenatal-integrated screening. In our models, the objective is to maximize the overall detection rate, and the constraints include upper-bounds on false positive rates based on age or age group: (1) *age-specific* risk-cutoff values for each individual age (A-IP), and (2) *age-group-specific* risk-cutoff values (AG-IP)to ensure that the same risk-cutoff value applies to women within an age group. The constraints enforce that the false positive rate for any age (or age group) does not exceed a selected maximum acceptable value. A-IP is more flexible, resulting in a specific cutoff value for each age, and is expected to lead to better outcomes; however, the large number of risk-cutoff values may be challenging for implementation in practice. In AG-IP,

there is a risk-cutoff value for each age group (rather than for each age), and hence, the number of risk-cutoff values is smaller and this policy provides a balance between the quality of outcomes and ease of implementation in practice.

4.2.1 Screening strategy and Monte Carlo simulations

Integrated screening is a commonly used prenatal test to assess the risk of Down syndrome in a pregnancy. The screening test is based on biomarkers of nuchal translucency, pregnancy associated plasma protein A (PAPP-A) in the first trimester as well as alpha-fetoprotain, human chorionic gonadotropin (hCG), unconjugated estriol, and Inhibin-A in the second trimester [11, 89]. The risk of a DS pregnancy is calculated using the biomarkers and the maternal age [88] and if the calculated risk exceeds a pre-defined risk-cutoff value, the result of the test is marked as positive. Women with a positive test result (i.e., a higher calculated risk than the risk-cutoff value) may consider amniocentesis. On the other hand, women whose test results are negative (i.e., the calculated risk is lower than the risk-cutoff value) often do not consider amniocentesis. Since the classification of the test result is not perfect, there are two potential adverse outcomes: (i) a false positive, followed by an amniocentesis possibly leading to a procedure-related fetal loss, and (ii) a false negative, leading to an undetected DS live birth. Detection rate refers to the fraction of DS pregnancies (also referred to as "affected" pregnancies) where DS is identified correctly by the test (note that the detection rate is equal to 1 minus the false negative rate).

We utilized a previous validated Monte Carlo simulation [93] to calculate the detection and false positive rates given a risk-cutoff value. In each run of Monte Carlo simulation, we sampled 800,000 affected (unaffected) pregnancies, applied risk-cutoff values, and calculated the detection rate (false positive rate) with respect to each risk-cutoff value [94]. In particular, the distributions of log10 biomarkers in affected and unaffected cohorts of pregnant women are modeled as truncated multivariate Gaussian

distributions with means, correlation matrices, and truncation limits reported in the Serum, Urine, and Ultrasound Screening Study (SURUSS) [90]. The uptake rate of amniocentesis, spontaneous fetal loss rate due to DS, and procedure-related fetal loss rate due to amniocentesis are set to 57.1%, 23%, and 0.3% respectively [93].

4.2.2 Integer programming model

We formulated two IP models, A-IP and AG-IP, with an objective function of maximizing the overall detection rate and constraints to ensure age-specific or age-group-specific, as well as overall upper bounds on false positive rates. The complete model formulations of A-IP and AG-IP can be found in §4.2.2.2. The models are solved using Gurobi (Gurobi Optimization, Inc. 2016).

4.2.2.1 Model inputs

To prepare the input parameters for the model, we utilized the Monte Carlo simulations described in §4.2.1. We generated an input table which contains a set of risk-cutoff values along with their corresponding detection and false positive rates for each maternal age (12 to 50 years old). A given risk-cutoff value results in a particular detection rate and false positive rate for a certain age. Hence, for the model to capture a wide range of possibilities in terms of (detection and false positive rates) combinations for different ages, it needs a carefully selected set of risk-cutoff values as input. We started with a basic set of risk-cutoff values ranging from 1/450 to 1/51 (with increments of 1 in the denominator). We then generated a set of risk-cutoff values corresponding to detection rates ranging from 50% to 99% (with increments of 1%) and false positive rates ranging from 1% to 40% (with increments of 0.5%) for each age. The union of these sets of potential risk-cutoff values forms the input for both A-IP and AG-IP. The corresponding steps to the above description for selecting the set of possible risk-cutoff values (to be the input into the model) are:

Step 1 **Select a basic set of risk-cutoff values**: We applied 400 cutoffs $(\frac{1}{51}, \frac{1}{52}, \cdots, \frac{1}{449}, \frac{1}{450})$

one by one to each age (12 to 50 years old) and calculated the detection rate and false positive rate for each age and risk-cutoff value pair.

Step 2 **Inversely calculate risk-cutoff values corresponding to a specific detection or false positive rate**: We used a bisection search to find out the risk-cutoff value that would result in a certain detection rate $(50\%, 51\%, \cdots, 99\%)$ or a false positive rate $(1\%, 1.5\%, 2\%, \cdots, 39.5\%, 40\%)$ for a given age. The feasibility of bisection search is guaranteed by the analytical result of Observation 1. The stopping criteria for our bisection search is the detection rate (or false positive rate) is within a 0.001% relative error rate of the target value of detection rate (or false positive rate). This step generates 129 risk-cutoff values for each age and guarantees that the risk-cutoff values in each age can cover the detection or false positive rates of interest.

Step 3 **Risk-cutoff values from Step 2 in a different age**: We applied the cutoffs generated from Step 2 for one age to each of other ages to make sure each age has the same set of risk-cutoff values. This step applies additionally 4902 $(= 129 \times 38)$ risk-cutoff values to each age.

A simplified version of Step 2: we randomly pick 30-years-old age as our base age and calculate the 129 risk-cutoff values using bisection searches as described in Step 2. Using the explicit formula in the condition of Proposition 1, we can directly map the risk-cutoff values calculated for the base age to those for the other 38 ages without any more inverse calculation, which is computationally intensive. On average, each bisection search takes 516 seconds using MATLAB parallel computing (two AMD Athlon II CPU cores running at 3.0 GHz with 6 GB of memory). In total, we can save more than 2.5×10^6 seconds using this simplified Step 2.

4.2.2.2 Model formulation

Sets and parameters used in A-IP (IP for age-specific risk-cutoff values) and AG-IP (IP for age-group-specific risk-cutoff values) are defined in Table 6. Each age group will be assigned one and only one risk-cutoff value, which is represented by decision variables in Table 6.

Table 6: Risk-cutoff value optimization using integer programming – sets, parameters, and decision variables.

Sets	
I	Set of maternal ages, $I = \{12, 13, \cdots, 49, 50\}$
R_i	Set of candidate risk-cutoff values for maternal age i, $i \in I$
I_k	Age group k, $\bigcup_{k=1,\ldots,K} I_k = I$, $I_k \bigcap I_{k'} = \emptyset, \forall k, k' \in \{1, 2, \cdots, K\}$
Parameters	
K	Total number of age groups considered in AG-IP
$DR_{i,r}$	Detection rate of using risk-cutoff value r in age i, $i \in I, r \in R_i$
$FPR_{i,r}$	False positive rate of using risk-cutoff value r in age i, $i \in I, r \in R_i$
γ_i	Upper-bound of false positive rate in age i, $i \in I$
θ	Upper-bound of overall false positive rate
P_i	Maternal age distribution in year 2014 [87], $i \in I$
Q_i^0	Age risk of Down syndrome [62], $i \in I$
ε	Upper-bound of relative error of risk-cutoff values within an age group
Decision Variables	
$x_{i,r}$	Binary decision variable, $x_{i,r} = \begin{cases} 1, & r \text{ is assigned to age } i \\ 0, & \text{otherwise} \end{cases}, i \in I, r \in R_i$

We presented two policies of age-specific risk-cutoff values and age-group-specific risk-cutoff values respectively. While the former provides more flexibility and better

outcomes, the latter aims to provide a balance between high performance and easiness of implementation.

Age-specific risk-cutoff values integer programming (A-IP)

The model formulation for A-IP is as follows:

$$\max_{x_{i,r}} \quad \frac{1}{\sum_{i\in I} P_i Q_i^0} \sum_{i\in I} P_i Q_i^0 \sum_{r\in R_i} DR_{i,r} x_{i,r} \tag{1}$$

$$\text{s.t.} \quad \sum_{r\in R_i} FPR_{i,r} x_{i,r} \le \gamma_i \quad \forall i \in I; \tag{2}$$

$$\frac{1}{\sum_{i\in I} P_i (1-Q_i^0)} \sum_{i\in I} P_i (1-Q_i^0) \sum_{r\in R_i} FPR_{i,r} x_{i,r} \le \theta; \tag{3}$$

$$\sum_{r\in R_i} x_{i,r} = 1 \quad \forall i \in I; \tag{4}$$

$$x_{i,r} \in \{0,1\} \quad \forall i \in I, r \in R_i. \tag{5}$$

Constraints (2) ensure that false positive rate of each age does not exceed the age-based upper bound, and constraint (3) ensures that the upper bound of overall false positive rate. Constraints (4) guarantee that each age is assigned a risk-cutoff value. The objective is to maximize the overall detection rate.

Age-group-specific risk-cutoff values integer programming (AG-IP)

We consider age-group-specific risk-cutoff values. In particular, the risk-cutoff value is a step function of age group. Therefore, we add two sets of constraints to make sure the risk-cutoff values within the same age group are close enough (with $\varepsilon = 10^{-5}$).

The model formulation for AG-IP would be shown as:

$$\max_{x_{i,r}} \quad (1)$$

$$\text{s.t.} \quad (2) - (5);$$

$$\sum_{r \in R_i} r x_{i,r} - \sum_{r \in R_j} r x_{j,r} \leq \varepsilon \sum_{r \in R_i} r x_{i,r} \quad \forall i, j \in I_k, k \in \{1, \cdots, K\}; \tag{6}$$

$$\sum_{r \in R_j} r x_{j,r} - \sum_{r \in R_i} r x_{i,r} \leq \varepsilon \sum_{r \in R_i} r x_{i,r} \quad \forall i, j \in I_k, k \in \{1, \cdots, K\}. \tag{7}$$

4.2.3 Examples: age-specific and age-group-specific risk-cutoff values

We proposed two policies: age-specific risk-cutoff values (corresponding to A-IP) and age-group-specific risk-cutoff values (corresponding to AG-IP). While the former provides more flexibility and better outcomes, the latter offers ease of implementation while still improving the outcomes compared to the one-size-fits-all risk-cutoff value currently used in practice. The following are examples of these policies with specific parameter settings:

- **A-IP Eexample**: This example A-IP policy ensures upper bounds on age-specific false positive rates as shown in Figure 10 and also ensures a 4% upper bound on the overall false positive rate across all ages.

- **AG-IP Example**: We divided the cohort into three age groups: ages 12–29, 30–35, and 36–50. This example AG-IP policy ensures upper bounds on false positive rates, namely, 5%, 10%, and 30% for each of the three age groups, respectively (Figure 10) and also ensures a 5% upper bound on the overall false positive rate.

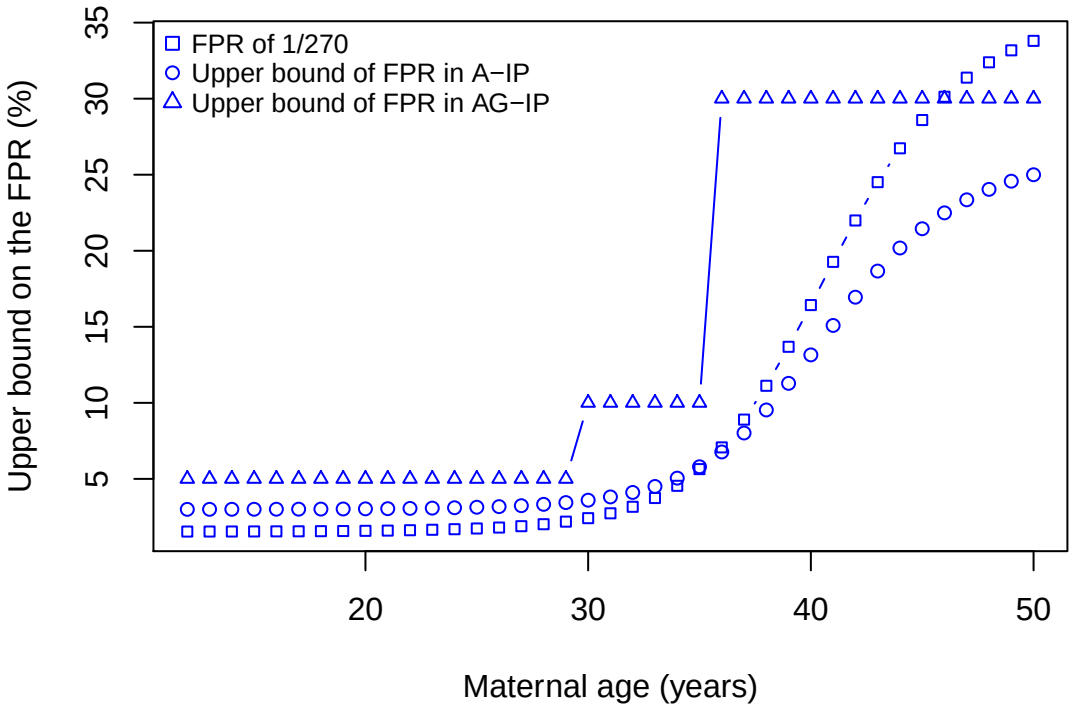

Figure 10: Age-specific and age-group specific upper bounds on FPR in Examples 1 and 2, used as input to solve the models A-IP and AG-IP. γ_i of A-IP is a linear transformation of FPR curve of 1/270. In particular, $\gamma_i = 3\% + (25\% - 3\%)(FPR_{i,1/270} - FPR_{12,1/270})/(FPR_{50,1/270} - FPR_{12,1/270})$. FPR under the one-size-fits-all risk-cutoff value of 1/270 is provided for comparison. FPR, false positive rate.

4.2.4 Performance measures

We reported the *detection rate* (DR) and *false positive rate* (FPR) for each age (or age group) resulting from the solutions, i.e., set of risk-cutoff values, of A-IP and AG-IP, applied to a cohort of 4 million simulated women. We also reported the numbers of (i) *undetected DS live births* (DSLs), as well as (ii) *euploid procedure-related fetal losses* (EFLs), corresponding to DR and FPR, respectively. We compared these results to the DR, FPR, DSL, and EFL resulting from the commonly used one-size-fits-all risk-cutoff value, 1/270.

4.3 Results

In Table 7, we compare the overall DR, FPR, EFL and DSL of the two examples with the one-size-fits-all risk-cutoff value of 1/270. The one-size-fits-all risk cutoff value of 1/270 results in DR=91.01% and FPR=3.56%. The age-specific risk-cutoff values in Example 1 result in DR=91.48% and FPR=4.00%. The age-group-specific risk-cutoff values in Example 2 result in DR=92.46% and FPR=4.99%), respectively, comparing to 12.21% under the risk-cutoff value of 1/270. In Examples 1 and 2, the overall DR is higher, and the DRs and FPRs in younger age groups 12–39 and 30–35 are also higher than those under 1/270. Compared to 1/270, in Examples 1 and 2 the EFLs are higher and the DSLs are lower, resulting in a lower total number of adverse outcomes. Figures 11–13 present age-(group-)specific risk-cutoff values of our two examples, and the age-specific FPR, DR resulting from the age-(group-)specific risk-cutoff values.

Table 7: Overall and age-group-dependent detection rate (DR), false positive rate (FPR), Down syndrome live births (DSLs), euploid procedure-related fetal losses (EFLs), and total number of adverse outcomes (Total) under the (i) one-size-fits all risk cutoff value 1/270, (ii) age-specific cutoffs (solutions of A-IP), and (iii) age-group-specific cutoffs (solutions of AG-IP).

Output metric		Risk-cutoff values		
		1/270	Age-specific cutoffs	Age-group-specific cutoffs
Overall	DR (%)	91.01	91.48	92.46
	FPR (%)	3.56	4.00	4.99
	DSLs	696	659	586
	EFLs	241	269	341
	Total	937	928	927
Ages 12–29	DR (%)	83.47	85.92	87.65
	FPR (%)	1.78	2.53	3.29
	DSLs	288	246	216
	EFLs	68	98	128
	Total	356	344	344
Ages 30–35	DR (%)	88.39	89.61	91.68
	FPR (%)	3.53	4.32	6.08
	DSLs	237	211	171
	EFLs	75	91	129
	Total	312	302	300
Ages 36–50	DR (%)	95.69	94.89	94.98
	FPR (%)	12.21	10.27	10.41
	DSLs	171	202	199
	EFLs	98	80	84
	Total	269	282	283

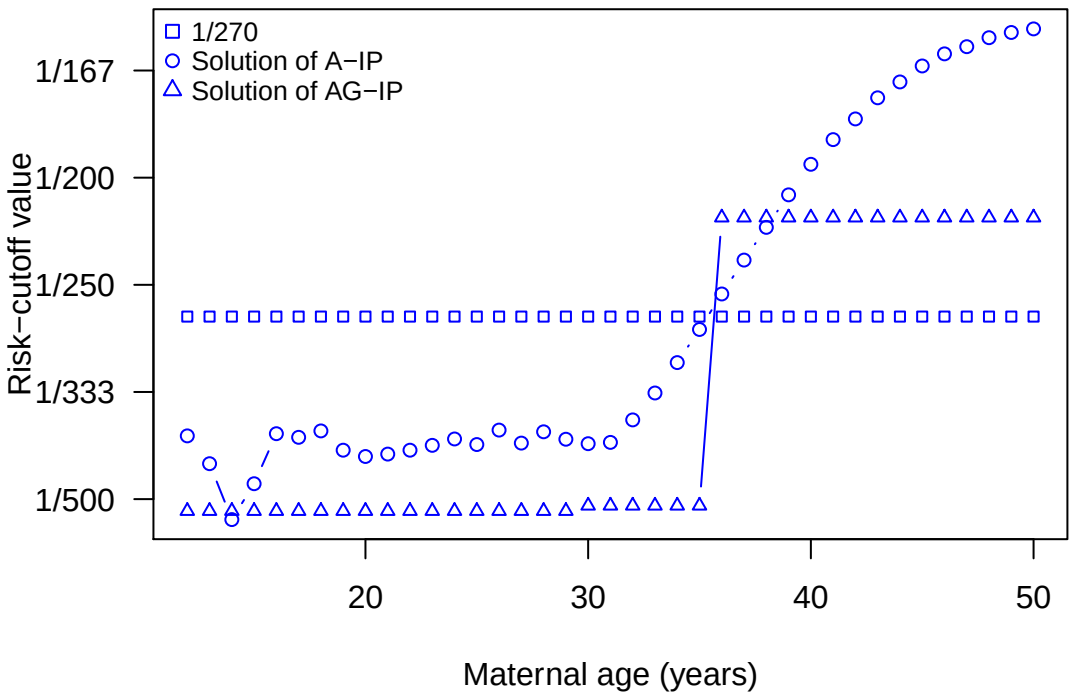

Figure 11: Solutions (risk-cutoff values) of age-specific risk-cutoff model (A-IP) and age-group-specific risk-cutoff model (AG-IP). The one-size-fits-all risk-cutoff value 1/270 is provided for comparison.

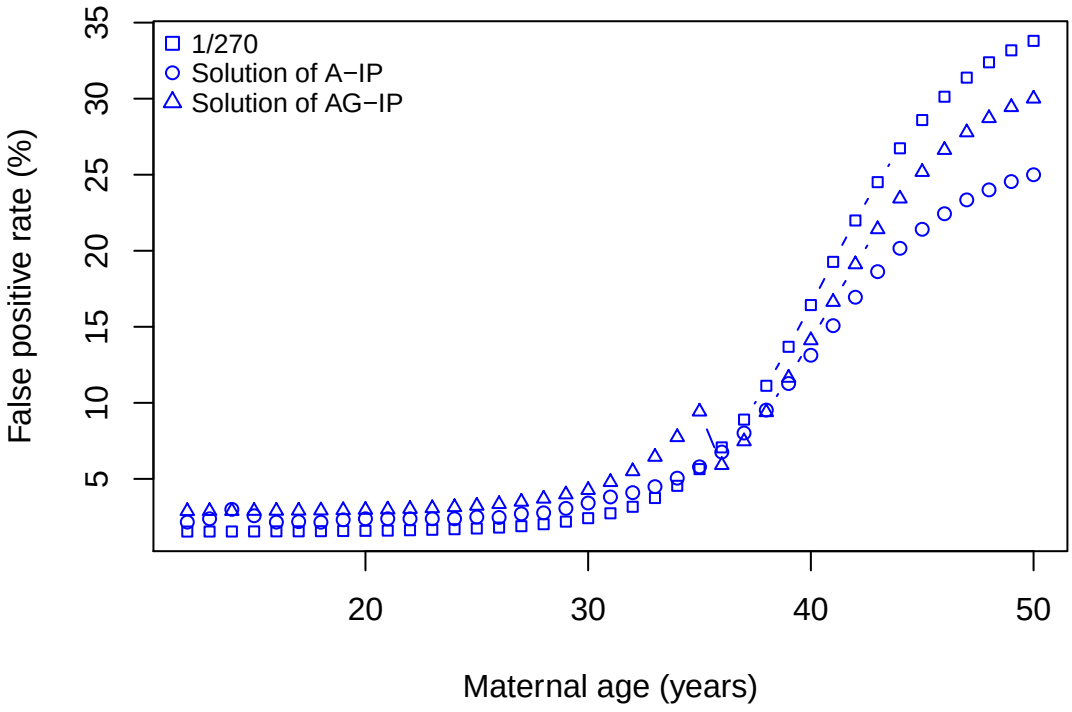

Figure 12: False positive rate resulting from the age-specific and age-group-specific risk-cutoff values in Examples 1 and 2 (output from the solutions of A-IP and AG-IP). False positive rate under the one-size-fits-all risk-cutoff value of 1/270 is provided for comparison.

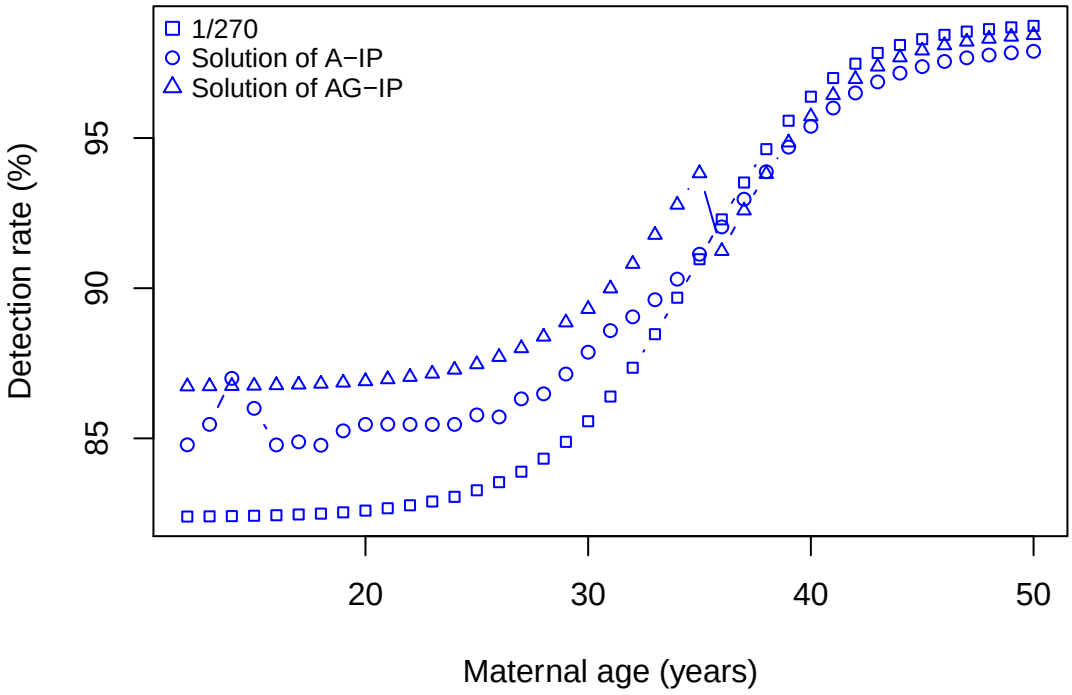

Figure 13: Detection rate resulting from the age-specific and age-group-specific risk-cutoff values in Examples 1 and 2 (output from the solutions of A-IP and AG-IP). Detection rate under the one-size-fits-all risk-cutoff value of 1/270 is provided for comparison.

4.4 Discussion

The two examples demonstrate that age-specific and age-group-specific risk-cutoff values have the potential to improve outcomes, e.g., by better controlling FPRs in older women without increasing the total number of overall adverse outcomes across all ages. The two examples are provided for illustration purposes only, whereas the actual choice of parameters for the potential implementation of age-specific or age-group-specific risk-cutoff values in practice would require a better understanding of the preferences of women across different age groups. The developed models are very flexible and provide the opportunity for what-if analysis and easy evaluation of

various parameter settings for age-specific and age-group-specific policies.

The results indicate a tradeoff (correlation) between DR and FPR. As shown in Table 7, age-specific and age-group-specific risk-cutoff values result in higher overall DR and higher overall FPR, whereas both the FPR and DR are lower for older women, compared to 1/270. To achieve a higher overall DR compared to that under 1/270, age-specific and age-group-specific risk-cutoff values slightly increase DR and FPR (resulting in lower DSLs and higher EFLs) in younger age groups. Hence, the optimal age-based risk-cutoff values balance the DR and FPR (and the corresponding adverse outcomes) among younger and older age groups.

4.5 Conclusion

Variation of FPRs across maternal ages resulting from a single one-size-fits-all risk-cutoff value has been highlighted in previous studies. We proposed two type of policies: age-specific and age-group-specific risk-cutoff values, and showed that those policies have the potential to mitigate the high variation of FPRs as well as to reduce the total number of adverse outcomes compared to commonly used risk-cutoff value of 1/270. The mathematical framework is very flexible to determine the age-(group-)specific risk-cutoff value based on the practical requirements.

4.6 Technical Appendix

We present the following findings in two observations and three propositions.

1. Detection (false positive) rate decreases with risk-cutoff value for a given age.

2. Detection (false positive) rate increases with age for a given risk-cutoff value.

3. Detection and false positive rate in two ages are equal when their risk-cutoff values satisfy an explicit formula.

4. We can always find a one-size-fits-all risk-cutoff value, which is optimal to minimize the positive-weighted sum of DSLs and EFLs.

5. We can always find a one-size-fits-all risk-cutoff value, which is optimal to maximize the overall detection rate with an upper bound constraint of the overall false positive rate.

Before getting into the proof, we introduce notations in addition to the ones in the previous sections. The post test risk given log10 biomarkers (denoted as $\mathbf{Z} \in R^6$) and maternal age i is denoted as $Q_i^1(\mathbf{Z})$ and it is calculated using standard method in [88], i.e., $Q_i^1(\mathbf{Z}) = \frac{\phi_1(\mathbf{Z})Q_i^0}{\phi_1(\mathbf{Z})Q_i^0+\phi_2(\mathbf{Z})(1-Q_i^0)(1-S)}$, where $\phi_1(\cdot)$ and $\phi_2(\cdot)$ are probability density functions of affected and unaffected pregnancies respectively, and S is a constant denoting the spontaneous fetal loss rate due to DS. Let r_i denote the risk-cutoff value for age i. $DR_{i,r_i}(FPR_{i,r_i})$ is the detection (false positive) rate for age i with risk-cutoff value r_i. By definition, detection (false positive) rate is the portion of affected (unaffected) pregnancies with post test risks higher than the risk-cutoff value, i.e., $DR_{i,x} = \int_{\mathbf{Z}:Q_i^1(\mathbf{Z})>x} \phi_1(\mathbf{Z})d\mathbf{Z}$ and $FPR_{i,x} = \int_{\mathbf{Z}:Q_i^1(\mathbf{Z})>x} \phi_2(\mathbf{Z})d\mathbf{Z}$.

Observation 1. *If $r_1, r_2 \in (0,1)$, and $r_1 \leq r_2$, then $DR_{i,r_1} \geq DR_{i,r_2}$ and $FPR_{i,r_1} \geq FPR_{i,r_2}$, $\forall i \in I$.*

Proof. Since $r_1 \leq r_2$, then $\{\mathbf{Z} : Q_i^1(\mathbf{Z}) > r_1\} \supseteq \{\mathbf{Z} : Q_i^1(\mathbf{Z}) > r_2\}$.

Therefore, it is obvious that for $\forall i \in I$,

$$DR_{i,r_1} = \int_{\mathbf{Z}:Q_i^1(\mathbf{Z})>r_1} \phi_1(\mathbf{Z})d\mathbf{Z} \geq \int_{\mathbf{Z}:Q_i^1(\mathbf{Z})>r_2} \phi_1(\mathbf{Z})d\mathbf{Z} = DR_{i,r_2}.$$

And similarly $FPR_{i,r_1} \geq FPR_{i,r_2}$, $\forall i \in I$. □

Observation 2. *If $i \leq j$, then $DR_{i,r} \leq DR_{j,r}$ and $FPR_{i,r} \leq FPR_{j,r}$, $\forall r \in (0,1)$.*

Proof. Consider the set of biomarkers of age i:

$$\left\{\mathbf{Z} : Q_i^1(\mathbf{Z}) > r\right\} = \left\{\mathbf{Z} : \frac{Q_i^1(\mathbf{Z})}{1-Q_i^1(\mathbf{Z})} > \frac{r}{1-r}\right\} = \left\{\mathbf{Z} : \frac{\phi_1(\mathbf{Z})}{\phi_2(\mathbf{Z})} > \frac{r(1-Q_i^0)(1-S)}{(1-r)Q_i^0}\right\}.$$

The order of ages, $i \leq j$, implies $Q_i^0 \leq Q_j^0$ and further $\frac{r(1-Q_i^0)(1-S)}{(1-r)Q_i^0} \geq \frac{r(1-Q_j^0)(1-S)}{(1-r)Q_j^0}$. Hence, $\{\mathbf{Z} : Q_i^1(\mathbf{Z}) > r\} \subseteq \{\mathbf{Z} : Q_j^1(\mathbf{Z}) > r\}$.

Therefore, $DR_{i,r} \leq DR_{j,r}$ and $FPR_{i,r} \leq FPR_{j,r}$. □

Proposition 1. *If* $r_i, r_j \in (0,1)$, *and* $\frac{\frac{r_i}{1-r_i}}{\frac{r_j}{1-r_j}} = \frac{\frac{Q_i^0}{1-Q_i^0}}{\frac{Q_j^0}{1-Q_j^0}}$, *then* $DR_{i,r_i} = DR_{j,r_j}$ *and* $FPR_{i,r_i} = FPR_{j,r_j}$, $\forall i, j \in I$.

Proof. Same as the proof above, we have

$$\left\{\mathbf{Z} : Q_i^1(\mathbf{Z}) > r_i\right\} = \left\{\mathbf{Z} : \frac{\phi_1(\mathbf{Z})}{\phi_2(\mathbf{Z})} > \frac{r_i(1-Q_i^0)(1-S)}{(1-r_i)Q_i^0}\right\}.$$

If $\frac{\frac{r_i}{1-r_i}}{\frac{r_j}{1-r_j}} = \frac{\frac{Q_i^0}{1-Q_i^0}}{\frac{Q_j^0}{1-Q_j^0}}$, then $\{\mathbf{Z} : Q_i^1(\mathbf{Z}) > r_i\} = \{\mathbf{Z} : Q_j^1(\mathbf{Z}) > r_j\}$.

Therefore, $DR_{i,r_i} = DR_{j,r_j}$ and $FPR_{i,r_i} = FPR_{j,r_j}$. □

Proposition 2. *There exists a one-size-fits-all type of risk-cutoff value which can minimize the positively weighted sum of the two adverse outcomes, i.e., for any* $W_1, W_2 > 0$, *the optimization problem*

$$\min_{[r_i]_{i\in I}} \quad W_1 \sum_{i\in I} NP_i Q_i^0 (1 - DR_{i,r_i})(1-S) + W_2 \sum_{i\in I} NP_i (1-Q_i^0) FPR_{i,r_i} DL \tag{8}$$

has an optimal solution as $r_{12} = r_{13} = \cdots = r_{50} = DL/(DL + (1-S)W_1/W_2)$, *where* N *denotes the total number of pregnant women in the cohort.*

Proof. Recall the definitions of DR_{i,r_i} and FPR_{i,r_i}:

$$1 - DR_{i,r_i} = \int_{\mathbf{Z}} \mathbb{1}(Q_i^1(\mathbf{Z}) \leq r_i)\phi_1(\mathbf{Z})d\mathbf{Z}; \tag{9}$$

$$FPR_{i,r_i} = \int_{\mathbf{Z}} \mathbb{1}(Q_i^1(\mathbf{Z}) > r_i)\phi_2(\mathbf{Z})d\mathbf{Z}. \tag{10}$$

Substituting the two equations to the objective function (8) and rearranging it,

we have

$$
\begin{aligned}
& N \sum_{i \in I} P_i \int_{\mathbf{Z}} \left[W_1 Q_i^0 \phi_1(\mathbf{Z}) \mathbb{1}(Q_i^1(\mathbf{Z}) \leq r_i)(1-S) + W_2(1-Q_i^0)\phi_2(\mathbf{Z}) \mathbb{1}(Q_i^1(\mathbf{Z}) > r_i) DL \right] d\mathbf{Z} \\
= & N \sum_{i \in I} P_i \int_{\mathbf{Z}} \left(Q_i^0 \phi_1(\mathbf{Z}) + (1-Q_i^0)\phi_2(\mathbf{Z}) \right) \left[W_1 Q_i^1(\mathbf{Z}) \mathbb{1}(Q_i^1(\mathbf{Z}) \leq r_i)(1-S) \right. \\
& \left. + W_2(1-Q_i^1(\mathbf{Z})) \mathbb{1}(Q_i^1(\mathbf{Z}) > r_i) DL \right] d\mathbf{Z} \\
= & N \sum_{i \in I} P_i \mathbb{E}_{\mathbf{Z} \sim pdf\left(Q_i^0 \phi_1(\mathbf{Z}) + (1-Q_i^0)\phi_2(\mathbf{Z}) \right)} \left[W_1 Q_i^1(\mathbf{Z}) \mathbb{1}(Q_i^1(\mathbf{Z}) \leq r_i)(1-S) \right. \\
& \left. + W_2(1-Q_i^1(\mathbf{Z})) \mathbb{1}(Q_i^1(\mathbf{Z}) > r_i) DL \right].
\end{aligned}
$$

The Monte Carlo simulation approximation of the expectation gives

$$
\begin{aligned}
& N \sum_{i \in I} P_i \frac{1}{N_i} \sum_{j=1}^{N_i} \left[W_1 Q_i^1(\mathbf{Z}_j) \mathbb{1}(Q_i^1(\mathbf{Z}_j) \leq r_i)(1-S) + W_2(1-Q_i^1(\mathbf{Z}_j)) \mathbb{1}(Q_i^1(\mathbf{Z}_j) > r_i) DL \right] \\
= & \sum_{i \in I} \sum_{j=1}^{N_i} \left[W_1 Q_i^1(\mathbf{Z}_j) \mathbb{1}(Q_i^1(\mathbf{Z}_j) \leq r_i)(1-S) + W_2(1-Q_i^1(\mathbf{Z}_j)) \mathbb{1}(Q_i^1(\mathbf{Z}_j) > r_i) DL \right].
\end{aligned}
$$

where $N_i = NP_i$ denotes the total number of women in age i, and $\mathbf{Z}_j$ follows pdf of $Q_i^0 \phi_1(\mathbf{Z}) + (1-Q_i^0)\phi_2(\mathbf{Z})$ for a given age i.

In other words, we can interpret the calculation of the two adverse outcomes in two equivalent ways: one uses prior probability and the other uses posterior probability.

- A cohort's point of view with age risk (prior probability): This is exactly (8). In the first term, $NP_i Q_i^0 (1 - DR_{i,r_i})(1-S)$, is the expected number of DS live births in maternal age i with risk-cutoff value r_i, and W_1 is the weight for an undetected DS live birth. Similarly, the number of euploid procedure-related fetal losses with r_i in age i can be calculated as $NP_i(1-Q_i^0)FPR_{i,r_i}DL$ with weight W_2.

- A woman's point of view with post test risk (posterior probability): A cohort consists of a large number of individuals and each individual woman (denoted as j) may contribute to the objective function in the following way:

$$
W_1 Q_i^1(\mathbf{Z}_j) \mathbb{1}(Q_i^1(\mathbf{Z}_j) \leq r_i)(1-S) + W_2(1-Q_i^1(\mathbf{Z}_j)) \mathbb{1}(Q_i^1(\mathbf{Z}_j) > r_i) DL. \quad (11)
$$

First we optimize the problem for each woman. If we can find a common optimal solution which is independent of log10 biomarkers $\mathbf{Z}_j$ and age i, then this solution will also be the optimal solution from the cohort's point of view.

Note that the optimal risk-cutoff value r^* for a given woman should satisfy:

$$r^* \leq p_1 \quad \forall p_1 \text{ satisfying } W_1 p_1 (1-S) \leq W_2 (1-p_1) DL; \tag{12}$$

$$r^* \geq p_2 \quad \forall p_2 \text{ satisfying } W_1 p_2 (1-S) > W_2 (1-p_2) DL. \tag{13}$$

It implies that $r^* = DL/(DL + (1-S)W_1/W_2)$, which can minimize (11) for each woman. Apparently, it is independent of $\mathbf{Z}_j$ and i, so it should also be the optimal solution for (8). □

Proposition 3. *There exists a one-size-fits-all type of risk-cutoff value which can maximize overall detection rate while keeping overall false positive rate no greater than an upper bound $\theta \in (0,1)$, i.e., the optimization problem*

$$\begin{aligned} \max_{[r_i]_{i\in I}} \quad & \frac{1}{\sum_{i\in I} P_i Q_i^0} \sum_{i\in I} P_i Q_i^0 DR_{i,r_i} \\ s.t. \quad & \frac{1}{\sum_{i\in I} P_i (1-Q_i^0)} \sum_{i\in I} P_i (1-Q_i^0) FPR_{i,r_i} \leq \theta. \end{aligned}$$

has an optimal solution the type of which is one-size-fits-all ($r_{12} = r_{13} = \cdots = r_{50}$).

Proof. Let $r^* = [r_i^*]_{i\in I}$ denote an optimal age-specific solution of the problem. Because of the monotonicity of detection and false positive rates as functions of the risk-cutoff value, the false positive rate of any optimal solution should attain the right-hand side, i.e.,

$$\frac{1}{\sum_{i\in I} P_i (1-Q_i^0)} \sum_{i\in I} P_i (1-Q_i^0) FPR_{i,r_i^*} = \theta.$$

On the other hand, we can use bisection search to find out the one-size-fits-all type of risk-cutoff value that make the constraint tight, and denote the value as r^0

(where $r_{12} = r_{13} = \cdots = r_{50} = r^0$). Thus, we have

$$\frac{1}{\sum_{i\in I} P_i(1-Q_i^0)} \sum_{i\in I} P_i(1-Q_i^0) FPR_{i,r^0} = \theta.$$

In addition, there must exist $W_1/W_2 > 0$ satisfying $r^0 = DL/(DL + (1-S)W_1/W_2)$.

Construct an unconstrained optimization problem:

$$\begin{aligned} \max_{[r_i]_{i\in I}} \quad F_1([r_i]_{i\in I}) =& W_1/W_2 \frac{(1-S)\sum_{i\in I} P_i Q_i^0}{DL \sum_{i\in I} P_i(1-Q_i^0)} \frac{\sum_{i\in I} P_i Q_i^0 DR_{i,r_i}}{\sum_{i\in I} P_i Q_i^0} \\ & - \frac{\sum_{i\in I} P_i(1-Q_i^0) FPR_{i,r_i}}{\sum_{i\in I} P_i(1-Q_i^0)}. \end{aligned} \tag{14}$$

Note that $W_1/W_2 \frac{(1-S)\sum_{i\in I} P_i Q_i^0}{DL \sum_{i\in I} P_i(1-Q_i^0)}$ is a positive constant.

After multiplying a negative constant and adding a constant term, (14) is equivalent to the following optimization problem:

$$\begin{aligned} \min_{[r_i]_{i\in I}} \quad F_2([r_i]_{i\in I}) =& W_1 \sum_{i\in I} NP_i Q_i^0 (1 - DR_{i,r_i})(1-S) \\ & + W_2 \sum_{i\in I} NP_i(1-Q_i^0) FPR_{i,r_i} DL. \end{aligned} \tag{15}$$

According to Proposition 2, there exists a one-size-fits-all type of risk-cutoff value $r_{12} = r_{13} = \cdots = r_{50} = r^0$ that is optimal to the optimization problem (15). This implies that

$$\begin{aligned} & F_2([r^0, r^0, \cdots, r^0]) \le F_2(r^*) \\ \Rightarrow & F_1([r^0, r^0, \cdots, r^0]) \ge F_1(r^*) \\ \Rightarrow & W_1/W_2 \frac{(1-S)\sum_{i\in I} P_i Q_i^0}{DL \sum_{i\in I} P_i(1-Q_i^0)} \frac{\sum_{i\in I} P_i Q_i^0 DR_{i,r^0}}{\sum_{i\in I} P_i Q_i^0} - \theta \\ & \ge W_1/W_2 \frac{(1-S)\sum_{i\in I} P_i Q_i^0}{DL \sum_{i\in I} P_i(1-Q_i^0)} \frac{\sum_{i\in I} P_i Q_i^0 DR_{i,r_i^*}}{\sum_{i\in I} P_i Q_i^0} - \theta \\ \Rightarrow & \frac{1}{\sum_{i\in I} P_i Q_i^0} \sum_{i\in I} P_i Q_i^0 DR_{i,r^0} \ge \frac{1}{\sum_{i\in I} P_i Q_i^0} \sum_{i\in I} P_i Q_i^0 DR_{i,r_i^*}, \end{aligned}$$

which suggests that the risk-cutoff value of r^0 is an optimal solution of the original optimization problem. □

CHAPTER V

CAPACITY ALLOCATION IN A SCHOOL-BASED ASTHMA CARE MODEL FOR PEDIATRIC PATIENTS

5.1 Introduction

Asthma is a chronic disease of the lungs and the most common childhood chronic illness in the United States [23], affecting about 6.3 million (8.6% of) American children [64]. In the state of Georgia, the asthma prevalence among children is even higher, i.e., 10.8% with a 95% confidence interval of 9.1–12.7% [44]. Childhood asthma is a leading cause of hospitalizations, emergency department visits, and missed school days [60]. In the United States, 128,000 hospital discharges of children under 15 years old had asthma as the first-listed diagnosis in 2010 [4]. During 2011, children under 15 years old had about 611,000 emergency department visits [22]. 13.8 million of school days were missed due to asthma, and 49% of asthma patients between 5 and 17 years old reported at least one missed school days during 2013 [24].

Although not completely curable, severe health outcomes of asthma, including hospitalizations, emergency department visits, as well as missed school days, can be reduced or even eliminated if the disease is well managed. In addition, with the proper treatment, asthma patients and their caregivers can keep the symptoms under control and prevent damage to lungs [60]. The reason for severe outcomes occurring every year mainly lies in the lack of access and adherence to appropriate asthma care. On the other hand, a significant portion of pediatric asthma patients have limited or no healthcare access while evidence suggests that lack of access is one of the leading causes for severe asthma outcomes. In particular, children living in inner cities are found to have an elevated risk of hospitalizations, emergency department visits, and

deaths due to asthma [23].

Several school-based programs have been established to reduce severe health outcomes and improve the quality of life among underserved children in several metropolitan areas in the United States, including Baltimore [15], Chicago [72], Los Angeles [52, 53], and St. Louis [86] among others. Those programs usually reach out to schools proactively via units of mobile asthma clinics staffed by a team of specialists, and work closely with school nurses to deliver care to underserved populations periodically (typically on a monthly basis). Such programs were shown to improve overall outcomes by improving access to care, increasing patient awareness and adherence to treatment. In particular, studies assessing the success of the school-based models have shown significant reductions in emergency room visits [15, 72], hospital days [15, 72], and missed work days of caregivers [72] as well as significant increases in symptom-free days and controller medication availability [15].

In the summer of year 2015, Children's Healthcare of Atlanta (Children's) Asthma Center at Hughes Spalding, a local hospital network in the state of Georgia, started to plan a school-based program with one unit of mobile care clinic for underserved pediatric asthma patients in metro Atlanta. On October 31, 2016, the first mobile asthma clinic was launched to serve metro Atlanta's pediatric patients [78].

The mobile asthma clinic is a 40-foot long, 8-foot wide vehicle built specifically for delivering pediatric health care services [76], and the program is funded by local charities [75]. The asthma services, including screening and diagnosis, medication prescriptions, and education, are free of charge to patients in this program [77]. The mobile asthma clinic includes two patient examination rooms, a draw station, 1 pulmonary function laboratory, and reception and waiting areas and is staffed by one pediatric nurse practitioner, two licensed practical or registered nurses, and one mobile outreach technician [78]. It runs four days a week and on average can take sixteen appointments each day.

In this paper, we focuses on building a mathematical framework to determine the capacity allocation by answering two-level questions: (i) which schools to visit, and (ii) if a school is visited, which groups of patients within the school to schedule for a limited number of appointments. The objective is to maximize the health outcomes measured in quality of life (QoL) months for the entire patient population. In particular, we propose a finite-horizon dynamic programming model and formulate a mixed-integer programming (MIP) for solving this model. In addition, we develop two heuristic algorithms, which can be easily implemented in practice. We parameterize our model using the data of a local public school district and extensive computational experiments show that the two heuristics are computationally-efficient and competitive in the quality of solutions.

5.2 Literature Review

There have been several studies on the disparities of pediatric asthma care within several states in the US. Garcia et al. [37] evaluate the geographic access to primary and asthma specialist care on severe outcomes, including emergency department visits and hospitalizations, of school-aged children (5–17 years old) in Georgia and North Carolina. They find the association between spacial access and severe outcomes is different across states. In particular, they show that access to primary care impact more than access to specialist care on Georgia emergency department visits, and the reverse applies for Georgia hospitalization. Fitzpatrick et al. [36] study the spatial access to asthma and compare the access to asthma care in 14 states in the United States. Bost et al. [16] compare the pediatric asthma service utilization, expenditures, and treatment with medications in Georgia, Florida, and North Carolina during 2006-2009 based on Medicaid claims data. In addition, the success and positive impacts on health outcomes achieved by mobile asthma clinics have been demonstrated in several observational studies [15, 52, 53, 72] with pre-post program comparisons.

Disease modeling has been an important component in the public health research to improve decision making. Ayer et al. [9] present a comprehensive survey of studies on disease screening, especially cancer screening, management of chronic diseases, as well as control of infectious diseases. They have pointed out the insufficiency and potential in research of chronic disease by applying operations research and management science techniques. Kucukyazici and Verter [54] review and highlight three studies on community-based care management for chronic diseases, including mental health care, post-stroke care, and asthma care. All those examples of disease management embedded Markov models of disease progression or disease-specific chronic care process. For disease modeling of asthma patients, Paltiel et al. [70] introduce a Markov state-transition simulation model named "Asthma Policy Model", which plays an important role in cost-effectiveness studies of various therapy options considering both clinical and economic outcomes.

Decision making in resource allocation consists of a broad range of research problems featured by limited budgets, funding, health care personnels, medical devices and equipments, vaccines or medicines, operation rooms, hospital beds, or appointments and so on. Griffin et al. [46] present a survey of resource allocation problems for infectious diseases such as malaria prevention. In addition, there are many recent studies on HIV-related resource allocation [33, 50], Hepatitis C treatment allocation [10] and organ transplant [8, 14]. Appointment scheduling problems in health care systems have been reviewed in Batun and Begen [12], Cayirli and Veral [21], Gupta and Denton [47]. Adaptive behaviors, such as no shows in the setting of appointment scheduling, have gained an increasing attention in recent literature [31].

However, there exist only few studies to combine clinical models of disease progression for asthma and operational constraints of the health system. Deo et al. [32] build a Markov decision process (MDP) framework considering capacity allocation

and disease progression to maximize quality adjusted life years of the patient population, and apply the MDP model to a school-based asthma management program in Chicago. On top of their model, Savelsbergh and Smilowitz [82] further optimize the appointment schedule with no-show behaviors by incorporating patients' time-of-day preferences. However, the decisions in the two studies are based on a pre-defined schedule of school visits so they optimized capacity allocation within one specific school assuming that the school would be visited periodically.

As a newly developed school-based asthma program in metro Atlanta, the selection of schools is a critical component of the decision. However, to the best of our knowledge, no asthma model has integrated clinical decisions with two-level operational decisions to allocate capacity among and within schools. In this study, we aim to fill this gap. In addition, different from Deo et al. [32], we aggregate asthma patients of the same illness level and consider the overall transition among illness states as moving from one period to another. Consequently, our model becomes deterministic and can be characterized as decisions at the tactical level. In addition, we formulate the problem as a mixed integer programming so that we are able to solve efficiently for large scale school systems with large number of patients, and quantify solutions as well as their performances, such as optimality gaps, which are extremely useful for practice use.

5.3 Mathematical Model

Our goal is to maximize the aggregate quality of life (QoL) months of the asthma patient population by making two-level decisions: (i) which schools to visit, and (ii) which groups of patients within the schools to schedule for a limited number of appointments. This problem can be viewed as a finite-horizon dynamic programming (DP) model which combines operational decisions and a clinic model of childhood asthma. To solve it, we formulate the capacity allocation problem as a mixed integer

programming (MIP).

5.3.1 Assumptions and notations

We limit the visit of the mobile asthma clinic to one school per day, and we allow to visit one school multiple times in a decision period, i.e., one month. We assume that treatment occurs in each appointment. Since the treatment effect can last for some time, we assume one patient can be scheduled for at most one appointment each month. In addition, due to the existence of surveys and portable monitoring devices, we assume a separate assessment mechanism from each appointment and we can reply on school nurses to monitor the true illness states of each patient so that we can know exactly the patients for whom we will schedule appointments. The sequence of events in a period is presented as follows:

1. calculation of QoLs based on the state at the beginning of a period,
2. capacity allocation for the current period,
3. appointments of patients in the period (with treatment),
4. natural disease progression of all patients.

The notations used in our model are presented below:

- t: index of decision period (month), $t = 1, \ldots, T$;
- k: school index, $k = 1, \ldots, K$;
- N_k: number of asthma patients in school k;
- B: capacity of appointments per day, $B = 16$ patients/day;
- C: number of service days in each month;
- i: index of illness degree, $i \in I = \{1, 2, \ldots\}$;

- $s_{t,k,i}$: number of patients of illness degree i in school k at the begin of period t, $s_{t,k,i} \in R_+$ and $\sum_{i \in I} s_{t,k,i} = N_k, \forall t = 1, \ldots, T$, $k = 1, \ldots, K$. In particular, $s_{1,k,i}$ denotes the initial number of patients of illness degree i in school k;

- b_i: quality of life (QoL) score of one patient of illness degree i;

- $R_t(S_t)$: the total QALYs of the patient population in period t with state $S_t = [[s_{t,k,i}]_{i \in I}]_{k=1,\ldots,K}$ at the beginning of period t;

- $x_{t,k}$: decision variable of number of days spending in school k in period t. $x_{t,k}$ can take any non-negative integers no greater than C, i.e. $x_{t,k} \in \{0, 1, 2, \ldots, C\}$;

- $y_{t,k,i}$: decision variable of number of patients in illness degree i getting scheduled for an appointment in school k in period t, $y_{t,k,i} \in \{0, 1, 2, \ldots, \lfloor s_{t,k,i} \rfloor\}$;

- Denote $a_t = [x_{t,k}, [y_{t,k,i}]_{i \in I}]_{k=1,\ldots,K}, \quad \forall t = 1, \ldots, T$;

- P: transition probability matrix for natural disease progression. In particular, let $P_{i,j}$ denote the transition probability from i to j, where $i, j \in I$;

- Q: transition probability matrix for treatment. In particular, $Q_{i,j}$ denotes the transition probability from i to j, where $i, j \in I$;

- $S^0_{k,i}$: real data, the initial number of patients of illness degree i in school k.

5.3.2 Dynamic programming (DP) formulation

State Space: The state of the system at time t is a vector of schools with sub-vectors: $[[s_{t,k,i}]_{i \in I}]_{k=1,\ldots,K}$, where the inner vector is a list of numbers of patients in different states, and the outer vector combines all K schools.

Action Space, A_t: Only one school can be served per day, i.e. C school visits

per month.

$$\sum_{k=1}^{K} x_{t,k} = C \qquad \forall t = 1, \ldots, T-1;$$

$$x_{t,k} \in \{0, 1, 2, \ldots, C\} \qquad \forall t = 1, \ldots, T-1, k = 1, \ldots, K;$$

$$\sum_{i \in I} y_{t,k,i} = B x_{t,k}, \qquad \forall t = 1, \ldots, T-1, k = 1, \ldots, K;$$

$$y_{t,k,i} \in [0, s_{t,k,i}], \qquad \forall i \in I, t = 1, \ldots, T-1, k = 1, \ldots, K.$$

Transition: If treatment is not in place for school k, the states evolve from t to $t+1$ based on the natural disease progression. On the other hand, if treatment is in place, the allocation of capacity within school k needs to be considered:

1. For those who have treatment (i.e., $y_{t,k,j}$), the transition has two parts: the treatment effect (Q) and the natural disease progression (P). Same as in [32], we also assume treatment takes effect immediately;

2. For those who have no treatment (i.e., $s_{t,k,j} - y_{t,k,j}$), the transition is solely based on the natural disease progression (P).

The relations between t and $t+1$ ($t = 1, \ldots, T-1$): $\forall k$,

$$s_{t+1,k,i}([s_{t,k,j}]_{j \in I}, [x_{t,k}, [y_{t,k,j}]_{j \in I}]) = \begin{cases} \sum_{j \in I} s_{t,k,j} P_{j,i}, & x_{t,k} = 0; \\ \sum_{j \in I} y_{t,k,j} \sum_{l \in I} Q_{j,l} P_{l,i} & \\ + \sum_{j \in I} (s_{t,k,j} - y_{t,k,j}) P_{j,i}, & x_{t,k} \geq 1. \end{cases}$$

Rewards: As mentioned above, the sequence of events is defined as: state at the beginning of a month $\rightarrow$ QoLs for the month $\rightarrow$ action/decision $\rightarrow$ (one month) $\rightarrow$ state transition.

$R_t([[s_{t,k,i}]_{i \in I}]_{k=1,\ldots,K}) = \sum_{k=1}^{K} \sum_{i \in I} b_i s_{t,k,i}$.

Optimality equations with value function $u_t(S_t)$: The objective is to maximize the total QoLs from period 1 to T. The optimality equations are:

$$u_t([[s_{t,k,i}]_{i\in I}]_{k=1,\dots,K}) = R_t([[s_{t,k,i}]_{i\in I}]_{k=1,\dots,K}) + \max_{a\in A_t} u_{t+1}(S_{t+1}([s_{t,k,i}]_{i\in I}, a)),$$
$$t = 1,\dots,T-1;$$

$$u_T([[s_{T,k,i}]_{i\in I}]_{k=1,\dots,K}) = R_T([[s_{T,k,i}]_{i\in I}]_{k=1,\dots,K}).$$

5.3.3 Mixed-integer programming (MIP) formulation

The typical method to solve a dynamic problem with a finite horizon is backward induction. However, for our problem, as shown in the computational study, the state space and action space are too large to enumerate. A more efficient way is to formulate it as a MIP and solve the MIP using solvers like Gurobi (Gurobi Optimization Inc., 2016).

MIP I: Base Model

$$\max_{x_{t,k},y_{t,k,i}} \quad \sum_{t=1}^{T}\sum_{k=1}^{K}\sum_{i\in I} b_i s_{t,k,i} \tag{16}$$

$$\text{s.t.} \quad s_{1,k,i} = S^0_{k,i}, \quad \forall k = 1,\dots,K, i \in I; \tag{17}$$

$$s_{t+1,k,i} = \sum_{j\in I}(s_{t,k,j} - y_{t,k,j})P_{j,i} + \sum_{j\in I} y_{t,k,j}\sum_{l\in I} Q_{j,l}P_{l,i},$$
$$\forall t = 1,\dots,T-1, k = 1,\dots,K, i \in I; \tag{18}$$

$$\sum_{k=1}^{K} x_{t,k} \le C, \quad \forall t = 1,\dots,T-1; \tag{19}$$

$$\sum_{i\in I} y_{t,k,i} \le B x_{t,k}, \quad \forall t = 1,\dots,T-1, k = 1,\dots,K; \tag{20}$$

$$y_{t,k,i} \le s_{t,k,i}, \quad \forall t = 1,\dots,T-1, k = 1,\dots,K, i \in I; \tag{21}$$

$$x_{t,k} \in \mathbb{Z}_+, \quad \forall t = 1,\dots,T-1, k = 1,\dots,K; \tag{22}$$

$$y_{t,k,i} \in \mathbb{Z}_+, \quad \forall t = 1,\dots,T-1, k = 1,\dots,K, i \in I. \tag{23}$$

The objective function (16) maximizes the aggregate QoL months. (17) defines

the initial states and (18) captures the state transition from period t to period $t+1$. The capacity constraints among and within schools are presented in (19) and (20) respectively. (21)–(23) define the range of the decision variables, where $\mathbb{Z}_+ = \{0, 1, 2, \ldots\}$ denotes non-negative integers.

5.4 Heuristics

As the number of schools and patients increase, the size of the MIP gets very large, which requires intensive computational resources, and thus poses a challenge to solve the problem to optimality. In particular, the memory size becomes an issue. Even for the 89 schools example in our computational study, it requires 150G memory (exhausted the memory limit of the servers we can have) to load the entire problem and solve it to optimality. It will become intractable if we extend the school list within Atlanta Public School district to all the eligible public schools of the existing asthma patients in Children's Healthcare of Atlanta – Hughes Spalding campus (roughly 142 schools based on the visit data in year 2014). Therefore, it is critical to seek for viable solutions. For that purpose, we develop two heuristics: one is an index-based heuristic algorithm and the other one is a relaxing version of the base MIP model.

5.4.1 Heuristic 1 – index-based greedy algorithm

The idea of the heuristic comes from the optimization problem below to maximize the QoLs one-step forward given the current state $[s_{t,k,i}]_{k=1,\ldots,K,i\in I}$ in period $t \in \{1, 2, \ldots, T-1\}$:

MIP II: One-Step Forward Optimization

$$\max_{x_{t,k},y_{t,k,i}} \quad \sum_{k=1}^{K}\sum_{i\in I} b_i s_{t+1,k,i} \tag{24}$$

$$\text{s.t.} \quad s_{t+1,k,i} = \sum_{j\in I}(s_{t,k,j} - y_{t,k,j})P_{j,i} + \sum_{j\in I} y_{t,k,j}\sum_{l\in I} Q_{j,l}P_{l,i},$$

$$\forall k = 1,\ldots,K, i \in I; \tag{25}$$

$$\sum_{k=1}^{K} x_{t,k} \le C; \tag{26}$$

$$\sum_{i\in I} y_{t,k,i} \le B x_{t,k}, \quad \forall k = 1,\ldots,K; \tag{27}$$

$$y_{t,k,i} \le s_{t,k,i}, \quad \forall k = 1,\ldots,K, i \in I; \tag{28}$$

$$x_{t,k}, y_{t,k,i} \text{ are non-negative integers.} \tag{29}$$

Substituting $s_{t+1,k,i}$ in (24) with (25), we have

$$\max_{x_{t,k},y_{t,k,i}} \quad \sum_{k=1}^{K}\sum_{i\in I} b_i \left(\sum_{j\in I}(s_{t,k,j} - y_{t,k,j})P_{j,i} + \sum_{j\in I} y_{t,k,j}\sum_{l\in I} Q_{j,l}P_{l,i}\right) \tag{30}$$

With (30), we can construct an index system for each illness degree.

Index for illness degrees: we rank the illness degree in a descending order of the marginal improvement of making an appointment for a patient of illness degree i:

$$\Phi(i) = \sum_{j\in I} b_j \sum_{l\in I} Q_{i,l}P_{l,j} - \sum_{j\in I} b_j P_{i,j} \tag{31}$$

Reorder set I to define $I_\Phi = \{I_{(1)}, I_{(2)}, \ldots\}$ given $\Phi(I_{(1)}) \ge \Phi(I_{(2)}) \ge \ldots$.

Index for schools and capacity allocation policy: for a given period $t \in \{1,\ldots,T-1\}$, we determine the capacity allocation variables of $x_{t,k}$ and $y_{t,k,i}$ by iteratively calculate the marginal improvement for each school k (denoted as $\Psi_m(k), k \in \{1,\ldots,K\}$) assuming we will allocate a capacity B of one more day (indexed by $m \in \{1,\ldots,C\}$) to that school. The $\tilde{(\cdot)}$ denotes intermediate variables, where $\tilde{B}$ dynamically captures the remaining daily capacity when we allocate one-day capacity to a given school, $\tilde{s}_{k,m,i}$ presents the number of patients of illness degree i in school

k without appointment yet when we allocate capacity day by day in a given month, and integer $\tilde{y}_{k,m,i}$ shows the allocation of one-day capacity to a given illness degree i in school k. The heuristic for scheduling school and patients in period t is presented as Algorithm 1.

Algorithm 1: Greedy heuristic for scheduling school and patients in period $t = 1, 2, \ldots, T-1$

Set $\tilde{s}_{k,m=1,i} = s_{t,k,i} \forall k = 1, \ldots, K, i \in I$, $x_{t,k} = 0, \forall k = 1, \ldots, K$ and $y_{t,k,i} = 0, \forall k = 1, \ldots, K, i \in I$;

for $m = 1, \ldots, C$ **do**

  for $k = 1, \ldots, K$ **do**

    Set $\tilde{B} = B, \Psi_m(k) = 0, \tilde{y}_{k,m,i} = 0, \forall i \in I$;

    for $i = I_{(1)}, I_{(2)}, \ldots$ **do**

      Define $\tilde{y}_{k,m,i} = \lfloor \min\{\tilde{s}_{k,m,i}, \tilde{B}\} \rfloor$ and save it for later use;

      Update $\Psi_m(k) = \Psi_m(k) + \Phi(i)\tilde{y}_{k,m,i}$;

      Update $\tilde{B} = \tilde{B} - \tilde{y}_{k,m,i}$;

      if $\tilde{B} < 1$ **then**

        Break: jump out of the for-loop of i

  Let $\tilde{k}_m = \arg\max_k \Psi_m(k)$. If there are ties, randomly pick a school k among the ties;

  Update decision variables for school $\tilde{k}_m$: $x_{t,\tilde{k}_m} = x_{t,\tilde{k}_m} + 1$ and $y_{t,\tilde{k}_m,i} = y_{t,\tilde{k}_m,i} + \tilde{y}_{\tilde{k}_m,m,i}, \forall i \in I$;

  Update number of patients without appointment **for** *schools* $k = 1, \ldots, K$ **do**

    if $k = \tilde{k}_m$ **then**

      $\tilde{s}_{k,m+1,i} = \tilde{s}_{k,m,i} - \tilde{y}_{k,m,i}, \forall i \in I$

    else

      $\tilde{s}_{k,m+1,i} = \tilde{s}_{k,m,i}, \forall i \in I$

Output $x_{t,k}$ and $y_{t,k,i}, \forall k = 1, \ldots, K, i \in I$;

Update

$$s_{t+1,k,i} = \sum_{j \in I}(s_{t,k,j} - y_{t,k,j})P_{j,i} + \sum_{j \in I} y_{t,k,j} \sum_{l \in I} Q_{j,l}P_{l,i}, \quad \forall k = 1, \ldots, K, i \in I.$$

5.4.2 Heuristic 2 – MIP relaxation algorithm

We propose a second heuristic by iteratively solving a series of MIPs relaxations. In particular, we relax the integer constraints for future periods. We define parameter D as the number of periods in which we impose integer constraints. In other words, the time window for integer variables is specified as the current period to the succeeding $(D-1)$ periods, in total D periods. As we move along decision periods, we fix the decisions in the past periods, only re-solve the problem for the future periods.

MIP III (for t^c and D): Heuristic at the current period $t^c, t^c = 1, \ldots, T-1$ with a time window D, resulting in the optimal solution denoted as $x_{t,k}^{*,t^c}, y_{t,k,i}^{*,t^c}$

$$\max \quad (16)$$

$$\text{s.t.} \quad (17) - (21);$$

$$x_{t,k} = x_{t,k}^{*,t^c-1}, \quad \forall t = 1, \ldots, t^c - 1, k = 1, \ldots, K, i \in I; \tag{32}$$

$$y_{t,k,i} = y_{t,k,i}^{*,t^c-1}, \quad \forall t = 1, \ldots, t^c - 1, k = 1, \ldots, K, i \in I; \tag{33}$$

$$x_{t,k} \in \{0, 1, \ldots\}, \quad \forall t = t^c, \ldots, t^c + D - 1, k = 1, \ldots, K; \tag{34}$$

$$y_{t,k,i} \in \{0, 1, \ldots\}, \quad \forall t = t^c, \ldots, t^c + D - 1, k = 1, \ldots, K, i \in I; \tag{35}$$

$$x_{t,k} \geq 0, \quad \forall t = t^c + D, \ldots, T - 1, k = 1, \ldots, K; \tag{36}$$

$$y_{t,k,i} \geq 0, \quad \forall t = t^c + D, \ldots, T - 1, k = 1, \ldots, K, i \in I. \tag{37}$$

(32)–(33) are integer constraints for the previous periods. Note that for the initial case, i.e., $t^c = 1$, the two sets of constraints are ignored from the formulation, and there is no need to define $x_{t,k}^{*,0}$ and $y_{t,k,i}^{*,0}$. Integer constraints apply for the current period, t^c, with a time window D, i.e. (34)–(35). In addition, we relax the integer constraints for following periods after $t^c + D$, i.e., (36)–(37).

Algorithm 2: Relaxing period by period algorithm for scheduling school and patients given a predefined time window D

Set $t^c = 1$;

Solve MIP III for optimal solution $x^{*,1}_{t,k}(\forall t = 1, \ldots, T-1, k = 1, \ldots, K)$ and $y^{*,1}_{t,k,i}(\forall t = 1, \ldots, T-1, k = 1, \ldots, K, i \in I)$;

for $t^c = 2, \ldots, T - D$ **do**

    Solve MIP III for optimal solution $x^{*,t^c}_{t,k}(\forall t = 1, \ldots, T-1, k = 1, \ldots, K)$ and $y^{*,t^c}_{t,k,i}(\forall t = 1, \ldots, T-1, k = 1, \ldots, K, i \in I)$;

Report the optimal solution for MIP III for $(T-1)$.

5.5 Computational Study

In this section, we present both long-term and short-term decisions. We consider Atlanta Public Schools (APS) district located in metro Atlanta. The APS school district consists of 89 public elementary, middle, and high schools covering kindergarten (KK) and 1st to 12th grades. In the first subsection, we report the input parameters to estimate the size of asthma patients and define the scenarios and metrics. Because of the curse of dimensionality, the model becomes intractable either when decision horizon T becomes too long or the number of schools K increases. Therefore, we present a long-term setting (i.e., $T = 120$ months) of one school, which is an aggregated version of the 89 schools, and separately we consider the capacity allocations for the 89-school system with a shorter decision horizon (i.e., $T = 6, 12, 18, 24$ months). We vary the decision horizons, capacity levels, as well as initial distributions to understand their impacts on the optimal solution. In addition, we report the optimality gap of our models as well as the performance of our heuristics.

5.5.1 Input parameters and metrics

There were in total 50669 students enrolled in KK to 12th grade in Fall 2016 [41]. Number of patients in each school is estimated and rounded to the nearest integer by combining the enrollment data [41] and the prevalences of asthma by age (14.1%, 9.7%, 10% for ages 5–9, 10–14, 15–17 years respectively [45]) in Georgia. Combining above inputs, the total number of patients in the 89 schools is estimated as 5975. The daily capacity of one unit of asthma care mobile is set to be $B = 16$ (half an hour per patient and eight hours per day) based on the estimation from Children's.

We utilize the definitions in [32] for illness degrees of $I = \{1, 2, \ldots, 16\}$, transition probability matrices (P and Q, see §5.7), and corresponding quality of life (QoL) scores. In particular, [32] defines illness degrees jointly by control level (controlled, improved, unchanged, and worsened) and severity level (mild intermittent, mild persistent, moderate persistent, and severe persistent). The detailed definitions are presented in Table 8.

Table 8: Definitions of illness degree level i and quality of life score (per month) from Deo et al. [32].

Definition of illness degrees		Control level			
		Controlled	Improved	Unchanged	Worsened
Severity level	Mild intermittent	$i = 1$	$i = 2$	$i = 3$	$i = 4$
	Mild persistent	$i = 5$	$i = 6$	$i = 7$	$i = 8$
	Moderate persistent	$i = 9$	$i = 10$	$i = 11$	$i = 12$
	Severe persistent	$i = 13$	$i = 14$	$i = 15$	$i = 16$
Quality of life score		0.95	0.87	0.80	0.73

The initial number of patients in each illness degree of each school is governed by an initial distribution. We define three initial distributions since we do not have such data from Children's and use them to generate instances in our computational study.

In §5.8, we generate one instance as an example using each initial distribution and report the size of patient by illness degree after aggregating schools.

- **Initial distribution 1 - optimistic version**: 80%, 10%, 5%, 5% of the patients start in control level of controlled, improved, unchanged, and worsened, respectively, and they are uniformly distributed among severity levels;
- **Initial distribution 2 - neutral version**: 10%, 40%, 40%, 10% of the patients start in control level of controlled, improved, unchanged, and worsened, respectively, and they are uniformly distributed among severity levels;
- **Initial distribution 3 - pessimistic version**: 5%, 5%, 10%, 80% of the patients start in control level of controlled, improved, unchanged, and worsened, respectively, and they are uniformly distributed among severity levels.

For the simplicity of presentation, we aggregate 16 illness degrees to three illness groups. The definition is as follows.

- **Illness group I (Best illness states)** including $i = 1, 5, 9, 13$;
- **Illness group II (Moderate illness states)** including $i = 2, 3, 6, 7, 10, 11, 14, 15$;
- **Illness group III (Worst illness states)** including $i = 4, 8, 12, 16$.

Our code base is in MATLAB (MATLAB R2016a, The MathWorks Inc.) with the optimization solver of Gurobi 6.5 [48]. The models run on servers with 8-core CPU of Intel Xeon and memory of 56–150G. For each instance of optimization problem, the solver runs up to two hours (controlled over all instance for a fair comparison) for optimal solutions with a warm start of the better solution of heuristic algorithms 1 and 2 (with $D = 1$). We choose $D = 1$ because we found diminishing improvements in objective values but large increases in computational time by setting $D \geq 2$. For technical details of the solutions, we define and compare the following metrics:

- δ: *Optimality gap* (defined as the relative difference between upper bound and optimal objective value, or equivalently, $\frac{z_{\mathrm{UB}}-z_{\mathrm{Opt}}}{z_{\mathrm{Opt}}}$, where z_{UB} is the best upper bound and z_{Opt} is the objective value of the optimal solution at the time the solver stops);

- τ_{H}: *Heuristic gap*, (defined as the relative difference between objective value of solution from the heuristic and optimal objective value, or equivalently, $\frac{z_{\mathrm{Opt}}-z_{\mathrm{H}}}{z_{\mathrm{Opt}}}$, where z_{H} is the objective value of the heuristic), i.e., improvement of optimal solution from that heuristic);

- Run time in seconds of the heuristic.

5.5.2 Capacity allocation of one school over a decision horizon of $T = 120$ months

Figures 14–15 compare the optimal solution and the Heuristic 1 solution side by side for each of the three initial distributions. While the Heuristic 1 allocates almost all capacity to the patients in the worst illness states, the optimal solution allocate substantial capacity to both patients in the worst illness states (defined as *treatment*) and patients in the moderate illness states (defined as *prevention*). As the system evolves with time, we observe that number of patients in each illness group has a huge difference between the two solutions: optimal solution ends up having a lot more patients in the best states as well as the worst states, and Heuristic 1 solution results in a majority of patients in the moderate states. The gap between Heuristic 1 and the optimal solution is quite small (0.40–1.45%, Table 9). Given the strong performance and the huge advantage of Heuristic 1 in the run time (less than 1 second, Table 9), it is promising to reply on Heuristic 1 solution to schedule patients in practice.

Table 10 quantifies the improvement in the patients' illness states with more capacity ($C = 16, 32, 48$, i.e., 1,2,3 units of vehicles respectively): more patients (74.62–84.33% for $C = 48$ versus 30.17–38.42% for $C = 16$) are in the best illness states and much fewer patients (2.83–7.07% for $C = 48$ versus 23.29–38.31% for $C = 16$) are in the worst illness states. Combining with a fixed cost of \$500,000 [76] to launch one unit of vehicle on the road, as well as an estimation of incremental quality-adjusted life years (QALYs), we can conduct cost-effectiveness analysis of the capacity size (Table 11). The cost-effectiveness analysis can inform Children's decision of expanding the current program.

In the domain of medical decision making, the QoL evaluation commonly introduces a discounting factor (an annual discount rate of 3% or equivalently a monthly rate of 0.25%) to consider time effect. In §5.9, we define a model formulation with a discounting factor and present the corresponding solutions.

Figure 14: Capacity allocation to illness groups of one school over $T = 120$ months.

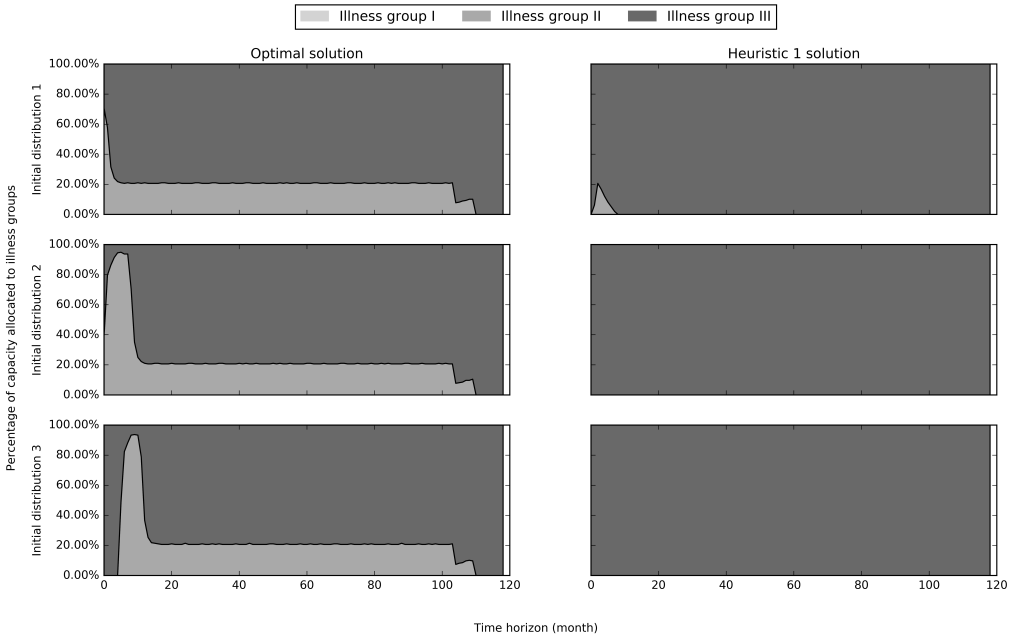

Figure 15: Number of patients in illness groups of one school over $T = 120$ months.

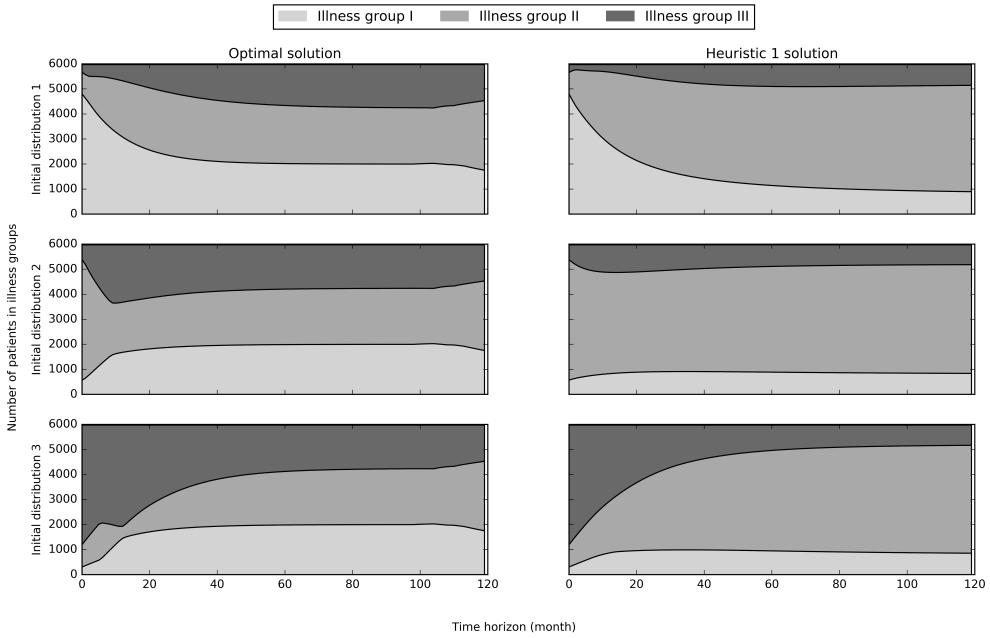

Table 9: Optimality gap and heuristic gaps of capacity allocation for one school with $T = 120$ months.

Metric	Initial distribution 1 (optimistic)		
	$C = 16$	$C = 32$	$C = 48$
δ	0.0022%	0.0050%	0.0072%
τ_{H1}	1.0598%	1.1508%	0.1250%
τ_{H2}	0.0000%	0.0000%	0.0000%
Run time (seconds) of heuristic 1	0.44	0.50	0.20
Run time (seconds) of heuristic 2	138.93	178.63	89.87
Metric	Initial distribution 2 (neutral)		
	$C = 16$	$C = 32$	$C = 48$
δ	0.0020%	0.0046%	0.0066%
τ_{H1}	1.4526%	1.1555%	0.4307%
τ_{H2}	0.0000%	0.0000%	0.0000%
Run time (seconds) of heuristic 1	0.44	0.53	0.19
Run time (seconds) of heuristic 2	132.77	181.94	86.69
Metric	Initial distribution 3 (pessimistic)		
	$C = 16$	$C = 32$	$C = 48$
δ	0.0019%	0.0042%	0.0064%
τ_{H1}	1.3238%	1.1007%	0.4022%
τ_{H2}	0.0000%	0.0000%	0.0000%
Run time (seconds) of heuristic 1	0.44	0.50	0.19
Run time (seconds) of heuristic 2	134.94	184.22	87.92

Table 10: Optimal capacity allocation and percentage of patients in the illness groups of one school over $T = 120$ months.

	Initial distribution 1 (optimistic)		
	$C = 16$	$C = 32$	$C = 48$
Capacity in illness group I (redundant)	0	0	0
Capacity in illness group II and III	30464	60928	91392
% capacity in illness group III out of those in illness groups II and III	80.48%	43.48%	21.75%
% patients in illness group I (best)	38.42%	63.51%	84.33%
% patients in illness group III (worst)	23.29%	3.78%	2.83%
	Initial distribution 2 (neutral)		
	$C = 16$	$C = 32$	$C = 48$
Capacity in illness group I (redundant)	0	0	0
Capacity in illness group II and III	30464	60928	91392
% capacity in illness group III out of those in illness groups II and III	76.49%	48.38%	25.04%
% patients in illness group I (best)	31.30%	54.69%	76.50%
% patients in illness group III (worst)	30.07%	5.20%	3.55%
	Initial distribution 3 (pessimistic)		
	$C = 16$	$C = 32$	$C = 48$
Capacity in illness group I (redundant)	0	0	0
Capacity in illness group II and III	30464	60928	91392
% capacity in illness group III out of those in illness groups II and III	78.39%	52.58%	28.84%
% patients in illness group I (best)	30.17%	52.73%	74.62%
% patients in illness group III (worst)	38.31%	10.63%	7.07%

Table 11: Incremental quality-adjusted life years (QALYs) over $T = 120$ months by increasing capacity.

	$\text{QALY}_{C=32} - \text{QALY}_{C=16}$	$\text{QALY}_{C=48} - \text{QALY}_{C=32}$
Initial distribution 1 (optimistic)	3057.95	1841.50
Initial distribution 2 (neutral)	3220.11	1974.87
Initial distribution 3 (pessimistic)	3299.59	2064.06

5.5.3 Capacity allocation of 89 schools over a decision horizon of up to $T = 24$ months

Figures 16–17 show the capacity allocation of optimal solution and Heuristic 1 solution over $T = 24$ months, and number of patients in illness groups as a result of the two solutions. Similar to the capacity allocation of one school over a long decision horizon, for the pessimistic initial distribution with $C = 16$ (one vehicle), Heuristic 1 allocates all capacity to treatment and results in a large portion of patients in the moderate illness states over time. The heuristic gap of Heuristic 1 is 0.10%, 0.22%, and 0.08% for initial distribution 1, 2, and 3 respectively (Table 12), which further suggests that Heuristic 1 has a strong performance. This is in line with the literature that such type of index-based myopic policy can perform very well [32, 79, 80, 81]. In addition, we find that the strong performance of Heuristic 1 persists in the sensitivity analyses of transition probability matrices P and Q, and QoL scores (§5.10). In addition, a setting with more schools (for example, in the school districts not limited to APS) or a decision horizon longer than $T = 24$ months will explode the memory size and easily make the MIP intractable. In that case, heuristics become the only option to solve for capacity allocation in those settings since they are showing a comparable solution quality to the MIP solutions (Table 12) when the MIP solutions are available.

Tables 13–14 quantify the portion of patients in the best and worst illness states as a result of the optimal solution of the scenarios with various capacity ($C =$

$16, 32, 48$, i.e., 1,2,3 vehicles respectively) and/or various decision horizons ($T = 6, 12, 18$ months). Not surprisingly, as capacity and/or decision horizon increase, more patients end up in the best illness states.

Figure 16: Capacity allocation to illness groups of 89 schools over $T = 24$ months.

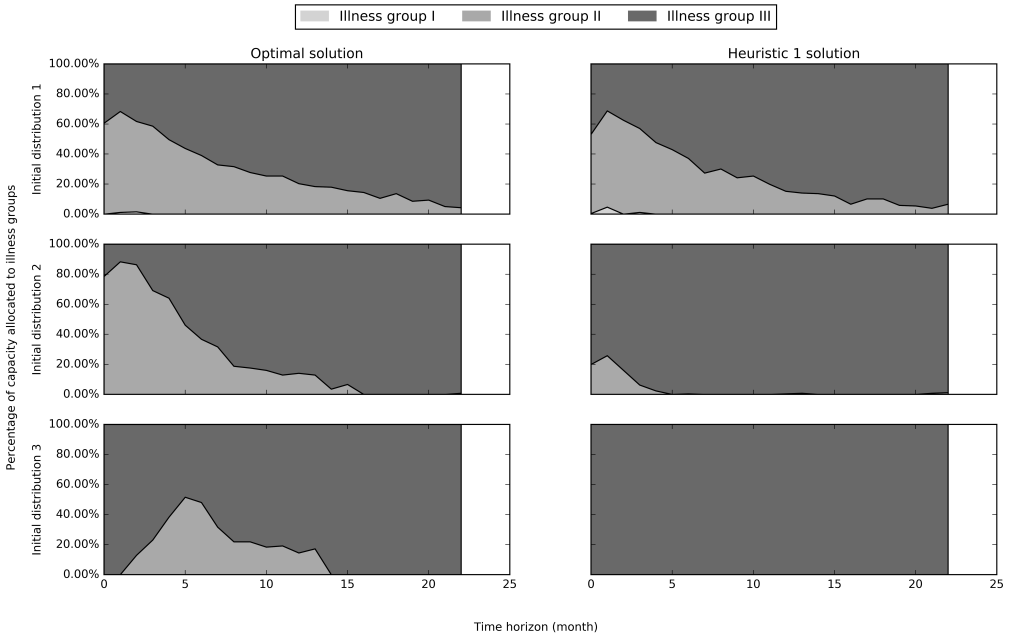

Figure 17: Number of patients in illness groups of 89 schools over $T = 24$ months.

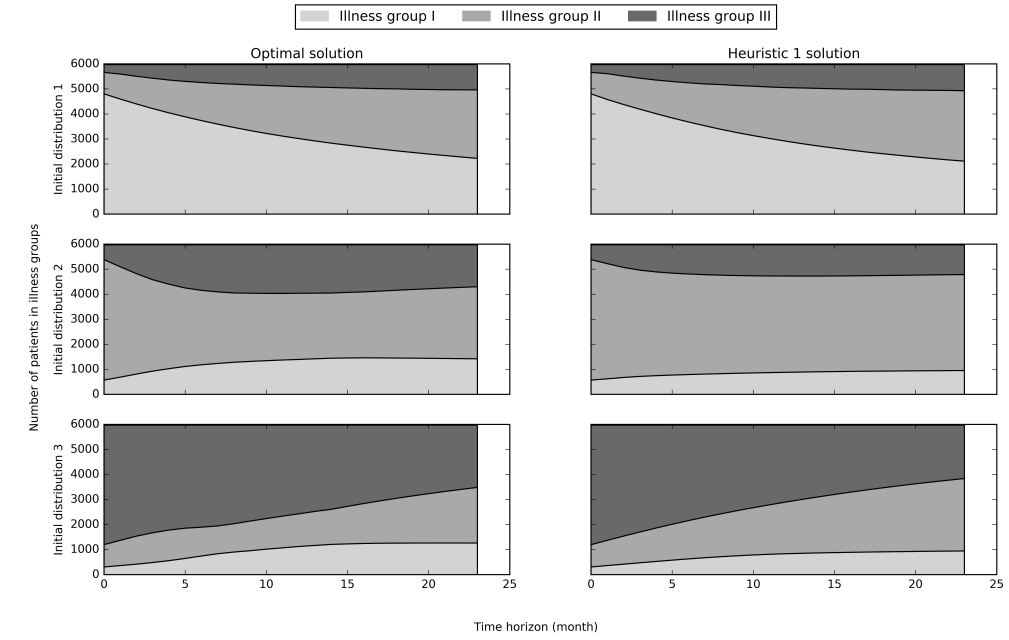

Table 12: Optimality gap and heuristic gaps of capacity allocation for 89 schools with $T = 24$ months.

Metric	Initial distribution 1 (optimistic)		
	$C = 16$	$C = 32$	$C = 48$
δ	0.2612%	0.6076%	1.2508%
τ_{H1}	0.1024%	0.0440%	0.0074%
τ_{H2}	0.0000%	0.0005%	0.0003%
Run time (seconds) of heuristic 1	0.38	0.42	0.44
Run time (seconds) of heuristic 2	1858.16	1859.81	2464.06
Metric	Initial distribution 2 (neutral)		
	$C = 16$	$C = 32$	$C = 48$
δ	0.1007%	0.3853%	0.5074%
τ_{H1}	0.2177%	0.3192%	0.3535%
τ_{H2}	0.0000%	0.0000%	0.0003%
Run time (seconds) of heuristic 1	0.39	0.40	0.43
Run time (seconds) of heuristic 2	1604.68	2015.57	2251.34
Metric	Initial distribution 3 (pessimistic)		
	$C = 16$	$C = 32$	$C = 48$
δ	0.0978%	0.2066%	0.3633%
τ_{H1}	0.0840%	0.1804%	0.1946%
τ_{H2}	0.0001%	0.0000%	0.0000%
Run time (seconds) of heuristic 1	0.42	0.39	0.46
Run time (seconds) of heuristic 2	1722.79	1405.68	2363.21

Table 13: Optimal capacity allocation and percentage of patients in the illness groups of 89 schools over $T = 24$ months.

	Initial distribution 1 (optimistic)		
	$C = 16$	$C = 32$	$C = 48$
Capacity in illness group I (redundant)	7	256	2803
Capacity in illness group II and III	5881	11520	14861
% capacity in illness group III out of those in illness groups II and III	71.26%	38.22%	26.86%
% patients in illness group I (best)	54.06%	66.19%	74.23%
% patients in illness group III (worst)	13.42%	9.16%	7.05%
	Initial distribution 2 (neutral)		
	$C = 16$	$C = 32$	$C = 48$
Capacity in illness group I (redundant)	0	3	85
Capacity in illness group II and III	5888	11773	17579
% capacity in illness group III out of those in illness groups II and III	73.74%	54.85%	35.33%
% patients in illness group I (best)	21.06%	33.63%	47.26%
% patients in illness group III (worst)	28.31%	13.10%	9.90%
	Initial distribution 3 (pessimistic)		
	$C = 16$	$C = 32$	$C = 48$
Capacity in illness group I (redundant)	0	1	34
Capacity in illness group II and III	5888	11775	17630
% capacity in illness group III out of those in illness groups II and III	86.16%	76.43%	54.71%
% patients in illness group I (best)	16.15%	27.62%	40.24%
% patients in illness group III (worst)	59.70%	36.17%	25.51%

Table 14: Optimal capacity allocation and percentage of patients in the illness groups of 89 schools for pessimistic initial distribution 3.

	Metric	$T = 6$	$T = 12$	$T = 18$
$C = 16$	Capacity in illness group I (redundant)	0	0	0
	Capacity in illness group II and III	1280	2816	4352
	% capacity in illness group III out of those in illness groups II and III	100.00%	100.00%	94.30%
	% patients in illness group I (best)	7.48%	9.86%	12.65%
	% patients in illness group III (worst)	72.95%	65.65%	61.62%
$C = 32$	Capacity in illness group I (redundant)	0	0	0
	Capacity in illness group II and III	2560	5632	8704
	% capacity in illness group III out of those in illness groups II and III	100.00%	99.11%	88.38%
	% patients in illness group I (best)	10.52%	15.13%	21.04%
	% patients in illness group III (worst)	63.87%	48.48%	40.76%
$C = 48$	Capacity in illness group I (redundant)	0	0	13
	Capacity in illness group II and III	3840	8448	13043
	% capacity in illness group III out of those in illness groups II and III	100.00%	81.75%	65.91%
	% patients in illness group I (best)	13.19%	22.43%	30.89%
	% patients in illness group III (worst)	54.92%	37.35%	28.68%

§5.11 presents how the capacity allocation to schools with various initial distributions and school sizes. In addition, §5.12 compares a variety of heuristic and policies. The optimal solution of 89 schools can still be very useful even we find that the Heuristic 1 solution has a very small gap (0.08% for the pessimistic initial distribution). By implementing the optimal solution instead of Heuristic 1, the absolute improvement

of objective value is about 96 QoL months for 5975 patients (on average of 0.8 week improvement per patient over 24 months), which is substantial in practice. We believe the improvement results from a better school selection and a better balance of treatment and prevention.

5.6 Conclusions

We develop a dynamic programming model to solve the capacity allocation problem for school-aged asthma patients within metro Atlanta. We formulate it as a MIP and propose two heuristics to solve it. We are the first to propose a model framework to consider school selection along with resources allocation within each selected school by taking disease progression into consideration. The numerical results show that the heuristics perform sufficiently well within a reasonable computational time, and can be useful especially when a larger number of schools or a longer decision horizon is considered and MIP is intractable due to the memory size. In addition to the policy of curing the sickest patients, controlling the disease progression among moderately sick patients is also important. The framework that we propose can be implemented in practice to determine the optimal schedule of the asthma program, and our findings are very useful for decision makers to plan for the school-based asthma program.

5.7 Appendix A: transition probability matrices

The transition probabilities matrices P and Q are based on Deo et al. [32] and we adjust some digits to make sure the summation of each row equals one (see Tables 15–16).

Table 15: The transition probability matrix P for natural disease progression based on Deo et al. [32].

	Illness degree															
i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0.97	0.01	0.01	0.01	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0.85	0.08	0.07	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0.93	0.04	0.01	0.02	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0.8	0.12	0.08	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0.93	0.07	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0.93	0.03	0.02	0.02	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0.79	0.11	0.1	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0.94	0.06	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0.9	0.04	0.03	0.03
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0.81	0.08	0.11
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.94	0.06
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Table 16: The transition probability matrix Q for treatment based on Deo et al. [32].

	Illness degree															
i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0.72	0.28	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0.7	0.21	0.09	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0.31	0.58	0.07	0.04	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0.59	0.41	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0.37	0.46	0.17	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0.24	0.67	0.06	0.03	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0.47	0.53	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0.35	0.45	0.2	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0.25	0.63	0.09	0.03	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0.57	0.43	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0.42	0.42	0.16	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0.35	0.53	0.09	0.03

5.8 Appendix B: one instance of initial number of patients corresponding to initial distributions 1–3

Table 17: One instance of initial patient size of each illness degree generated from initial distributions 1 (optimistic), 2 (neutral), 3 (pessimistic) respectively.

Illness degree i	Initial distribution 1	Initial distribution 2	Initial distribution 3
1	1184	128	70
2	129	606	58
3	80	579	139
4	75	155	1201
5	1195	148	90
6	143	567	58
7	64	623	161
8	85	149	1178
9	1227	153	77
10	146	620	76
11	72	600	176
12	84	156	1200
13	1199	145	72
14	161	600	73
15	63	615	150
16	68	131	1196

5.9 Appendix C: capacity allocation of one school over a decision period of $T = 120$ months with a discounting factor

We introduce a discounting factor and aggregate the QoLs-to-go after 120 months. In particular, we substitute the objective function of MIP I by adding a discounting factor for the quality of life scores over time, i.e., a discounting factor as $\lambda = \frac{1}{1+\gamma}$ with a monthly discount rate of $\gamma = 0.25\%$. In addition, we adjust the rewards in the ending period to combine all the values as if the illness state remains for infinite horizons afterwards. The objective function is updated from (16) to $\sum_{t=1}^{T-1} \lambda^{t-1} \sum_{k=1}^{K} \sum_{i \in I} b_i s_{t,k,i} + \frac{\lambda^{T-1}}{1-\lambda} \sum_{k=1}^{K} \sum_{i \in I} b_i s_{T,k,i}$. The constraints and decision

variables are the same as those in MIP I, i.e., constraints of (17)–(23) with decision variables of $x_{t,k}, y_{t,k,i}, \forall t = 1, \ldots, T-1, k = 1, \ldots, K, i \in I$. Table 18 shows a gap of 1.7–2.1% between Heuristic 1 and the optimal solution.

Table 18: Optimality gap and heuristic gaps of capacity allocation for one school with $T = 120$ months and a discount rate of $\gamma = 0.25\%$.

Metric	Initial distribution 1 (optimistic)		
	$C = 16$	$C = 32$	$C = 48$
δ	0.0025%	0.0047%	0.0083%
τ_{H1}	1.8991%	1.8328%	0.1695%
τ_{H2}	0.0000%	0.0002%	0.0000%
Run time (seconds) of heuristic 1	0.16	0.19	0.25
Run time (seconds) of heuristic 2	74.63	99.44	75.14
Metric	Initial distribution 2 (neutral)		
	$C = 16$	$C = 32$	$C = 48$
δ	0.0022%	0.0046%	0.0084%
τ_{H1}	2.0972%	1.8377%	0.2781%
τ_{H2}	0.0000%	0.0000%	0.0011%
Run time (seconds) of heuristic 1	0.32	0.25	0.26
Run time (seconds) of heuristic 2	106.68	102.99	71.52
Metric	Initial distribution 3 (pessimistic)		
	$C = 16$	$C = 32$	$C = 48$
δ	0.0023%	0.0045%	0.0098%
τ_{H1}	2.0383%	1.8216%	0.2711%
τ_{H2}	0.0000%	0.0001%	0.0000%
Run time (seconds) of heuristic 1	0.32	0.41	0.25
Run time (seconds) of heuristic 2	105.60	150.99	70.70

5.10 Appendix D: sensitivity analysis of transition probability matrices P and Q, and QoL scores for capacity allocation of 89 schools over a decision period of $T = 24$ months with one vehicle $C = 16$

We conduct probabilistic sensitivity analysis for transition probability matrices P and Q independently (Algorithm 2 in [95]). Table 19 shows that the gap between Heuristic 1 and the optimal solution is consistently small even with different P and

Q. It suggests that Heuristic 1 is able to capture the dynamics of the system very well and it is a strong candidate to be used in the capacity allocation problem when the MIP solution is not available.

Table 19: Range of Heuristic 1 gap in probabilistic sensitivity analysis of P and Q (10 instances along with the baseline).

	Heuristic 1 gap (τ_{H1})		
	mean	min	max
Initial distribution 1 (optimistic)	0.09%	0.05%	0.17%
Initial distribution 2 (neutral)	0.20%	0.09%	0.34%
Initial distribution 3 (pessimistic)	0.10%	0.01%	0.24%

The same as in [32], we also consider the concave and convex version of QoL scores:

- Concave QoL scores: 0.95, 0.90, 0.84, 0.73 for controlled, improved, unchanged, worsened illness degrees respectively;
- Linear QoL scores (used as the baseline scenarios in the main text): 0.95, 0.87, 0.80, 0.73 for controlled, improved, unchanged, worsened illness degrees respectively;
- Convex QoL scores: 0.95, 0.82, 0.76, 0.73 for controlled, improved, unchanged, worsened illness degrees respectively.

Table 20: Heuristic 1 gap for concave and convex QoL scores.

	Concave QoL scores	Convex QoL scores
Initial distribution 1 (optimistic)	0.06%	0.11%
Initial distribution 2 (neutral)	0.04%	0.36%
Initial distribution 3 (pessimistic)	0.04%	0.37%

It is true that under the current settings (linear QoL scores) the Heuristic 1 prioritizes treatment over prevention. However, in a more general setting, this might not be the case. In fact, in the convex QoL scores cases, Heuristic 1 is no longer a treatment-prioritized solution. In addition, Heuristic 1 considers the number of patients by illness degrees for each school and compare the total improvement among schools. Therefore, Heuristic 1 is not equivalent to a treatment-prioritized policy and the strong performance of Heuristic 1 does not imply that treatment-prioritized policy has a good performance as well. We believe that the small gap between Heuristic 1 and optimal solution is not because Heuristic 1 is a treatment-prioritized policy but because Heuristic 1 can well capture the main dynamics of the illness progression and respond to the current system states in each decision period.

5.11 Appendix E: capacity allocation by patient size and initial distribution for 89 schools over a decision period of $T = 24$ months with one vehicle $C = 16$

We group schools by their patient sizes (each group has a similar number of schools): number of patients between 7 and 52 (*small size schools*), 53 and 72 (*medium size schools*), 73 and 182 (*large size schools*), and uniformly assign initial distributions 1,2,3 to the 89 schools and test for 15 random instances. Figure 18 shows the average percentage of capacity (restricted to those for illness groups II and III, i.e., non-redundant capacity) by school size and initial distribution. We observe that the capacity allocation is driven by schools with a large size of patients and initial distribution 3.

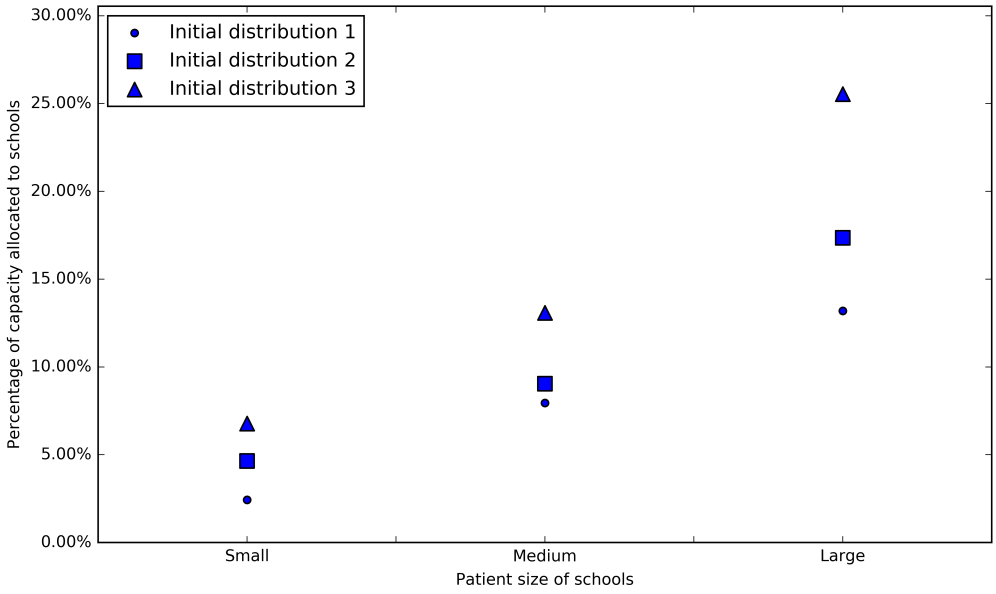

Figure 18: Average percentage of capacity allocation by patient size and initial distribution.

5.12 Appendix F: comparison of various policies for capacity allocation in 89 schools over a decision period of $T = 24$ months with one vehicle $C = 16$

We compare the objective value of each of the following policies and the relative gaps compared to the optimal capacity allocation for 89 schools (Table 21).

- Heuristic 1: index-based heuristic algorithm (Algorithm 1);
- *Do nothing*: the policy of not scheduling any patients (the baseline in [32]);
- *Treatment maximization*: a policy of maximizing the total number of patients in the worst illness states being scheduled (the optimal solution of a MIP with an objective function of maximizing treatment);
- *Treatment-prioritized heuristic*: a myopic policy of prioritizing schools based on the number of patients in the worst illness states and scheduling only patients

in the worst illness states of the prioritized schools (school and patient selection can change over time);

- *Prevention maximization*: a policy of maximizing the total number of patients in the moderate illness states being scheduled (the optimal solution of a MIP with an objective function of maximizing prevention);

- *Prevention-prioritized heuristic*: a myopic policy of prioritizing schools based on the number of patients in the moderate illness states and scheduling only patients in the moderate illness states of the prioritized schools (school and patient selection can change over time);

- *Fixed schedule policy*: a myopic policy of prioritizing schools based on the initial number of patients in the worst illness states and randomly scheduling patients in each period (the list of schools and days per month of visiting each school are fixed at the beginning of the entire decision horizon, and the schedule repeats every month);

- *One aggregating school policy*: optimal solution of an one school setting (relaxation of base MIP by removing school boundaries).

The policy of doing nothing provides a lower bound of the objective value (z_{Null}) of any policy. Compared to that, the fixed schedule policy can only close about 20–50% ($\frac{z_{\text{Fixed}}-z_{\text{Null}}}{z_{\text{Opt}}-z_{\text{Null}}}$) of the gap to optimal solution in terms of the objective value, although it is the easiest policy to implement. The Heuristic 1 consistently outperforms treatment-/prevention-prioritized heuristics. Compared to the policy of doing nothing, our Heuristic 1 can close about 95–99% ($\frac{z_{\text{H1}}-z_{\text{Null}}}{z_{\text{Opt}}-z_{\text{Null}}}$) of the policy gap.

Table 21: Objective value and policy gap of various policy with $C = 16$ (1 vehicle) and $T = 24$ months.

Policies	Metric	Initial distribution 1	Initial distribution 2	Initial distribution 3
Optimal capacity allocation of 89 schools	Objective value	126334.32	118009.87	113187.56
Heuristic 1	Objective value	126205.00	117752.91	113092.43
	Policy gap	0.10%	0.22%	0.08%
Do nothing	Objective value	121074.22	111990.39	106645.19
	Policy gap	4.16%	5.10%	5.78%
Treatment maximization	Objective value	124889.98	117279.94	112492.13
	Policy gap	1.14%	0.62%	0.61%
Treatment-prioritized heuristic	Objective value	124722.43	116409.81	112074.62
	Policy gap	1.28%	1.36%	0.98%
Prevention maximization	Objective value	122856.85	115078.84	108752.92
	Policy gap	2.75%	2.48%	3.92%
Prevention-prioritized heuristic	Objective value	123522.65	115160.75	107489.19
	Policy gap	2.23%	2.41%	5.03%
Fixed schedule policy	Objective value	122642.05	114875.76	108044.24
	Policy gap	2.92%	2.66%	4.54%
One aggregating school policy	Objective value	126930.26	118162.00	113331.62
	Policy gap	-0.47%	-0.13%	-0.13%

CHAPTER VI

CONCLUSION

This thesis has applied Operations Research techniques in two healthcare applications, i.e., DS screening and asthma care for school-aged patients. We have shown the great potential to improve patients outcomes by using our optimization frameworks. Our findings in this thesis can help patients, policy makers, and health care providers to make informed decisions.

REFERENCES

[1] ACOG, "ACOG practice bulletin number 77: screening for fetal chromosomal abnormalities," *Obstetrics & Gynecology*, vol. 109, pp. 217–227, 2007.

[2] ACOG, "Committee opinion number 545," *Obstetrics & Gynecology*, vol. 120, pp. 1532–1534, 2012.

[3] ACOG, "Committee opinion number 640: cell-free DNA screening for fetal aneuploidy," *Obstetrics & Gynecology*, vol. 126, no. 3, pp. e31–e37, 2015.

[4] American Lung Association, "Trends in asthma morbidity and mortality. table 16." `http://www.lung.org/assets/documents/research/asthma-trend-report.pdf`, 2012. Accessed: May 28, 2016.

[5] American Lung Association, "Asthma & children fact sheet." `http://www.lung.org/lung-health-and-diseases/lung-disease-lookup/asthma/learn-about-asthma/asthma-children-facts-sheet.html`, 2014. Accessed: May 28, 2016.

[6] American Lung Association, "Trends in asthma morbidity and mortality. table 16." `http://www.lung.org/assets/documents/research/asthma-trend-report.pdf`, last updated: September, 2012. Accessed: May 28, 2016.

[7] American Lung Association, "Asthma & children fact sheet." `http://www.lung.org/lung-health-and-diseases/lung-disease-lookup/asthma/learn-about-asthma/asthma-children-facts-sheet.html`, last updated: September, 2014. Accessed: May 28, 2016.

[8] Ata, B., Skaro, A., and Tayur, S., "Organjet: Overcoming geographical disparities in access to deceased donor kidneys in the united states," *Management Science*, 2016.

[9] Ayer, T., Keskinocak, P., and Swann, J., "Research in public health for efficient, effective, and equitable outcomes," in *INFORMS Tutorials in Operations Research: Bridging Data and Decisions*, pp. 216–239, INFORMS, 2014.

[10] Ayer, T., Zhang, C., Bonifonte, A., Spaulding, A., and Chhatwal, J., "Prioritizing hepatitis c treatment in u.s. prisons." Working paper, Georgia Institute of Technology, 2016.

[11] Ball, R. H., Caughey, A. B., Malone, F. D., Nyberg, D. A., Comstock, C. H., Saade, G. R., Berkowitz, R. L., Gross, S. J., Dugoff,

L., Craigo, S. D., Timor-Tritsch, I. E., Carr, S. R., Wolfe, H. M., Emig, D., and D'Alton, M. E., "First-and second-trimester evaluation of risk for down syndrome," *Obstetrics & Gynecology*, vol. 110, no. 1, pp. 10–17, 2007.

[12] Batun, S. and Begen, M. A., "Optimization in healthcare delivery modeling: methods and applications," in *Handbook of Healthcare Operations Management*, pp. 75–119, Springer, 2013.

[13] Benn, P. A. and Collins, R., "Evaluation of effect of analytical imprecision in maternal serum screening for down's syndrome," *Annals of Clinical Biochemistry*, vol. 38, no. 1, pp. 28–36, 2001.

[14] Bertsimas, D., Farias, V. F., and Trichakis, N., "Fairness, efficiency, and flexibility in organ allocation for kidney transplantation," *Operations Research*, vol. 61, no. 1, pp. 73–87, 2013.

[15] Bollinger, M. E., Morphew, T., and Mullins, C. D., "The breathmobile program: a good investment for underserved children with asthma," *Annals of Allergy, Asthma & Immunology*, vol. 105, no. 4, pp. 274–281, 2010.

[16] Bost, J., Johnson, K., Serban, N., and Swann, J., "Pediatric asthma baseline metrics for georgia, florida, and north carolina (2006-2009)." Working paper, Georgia Institute of Technology, 2015.

[17] Caughey, A. B., "Cost-effectiveness analysis of prenatal diagnosis: methodological issues and concerns," *Gynecologic and Obstetric Investigation*, vol. 60, no. 1, pp. 11–18, 2005.

[18] Caughey, A. B., Hopkins, L. M., and Norton, M. E., "Chorionic villus sampling compared with amniocentesis and the difference in the rate of pregnancy loss," *Obstetrics & Gynecology*, vol. 108, no. 3, pp. 612–616, 2006.

[19] Caughey, A. B., Kaimal, A. J., and Odibo, A. O., "Cost-effectiveness of down syndrome screening paradigms," *Clinics in Laboratory Medicine*, vol. 30, no. 3, pp. 629–642, 2010.

[20] Caughey, A. B., Kuppermann, M., Norton, M. E., and Washington, A. E., "Nuchal translucency and first trimester biochemical markers for down syndrome screening: a cost-effectiveness analysis," *American Journal of Obstetrics and Gynecology*, vol. 187, no. 5, pp. 1239–1245, 2002.

[21] Cayirli, T. and Veral, E., "Outpatient scheduling in health care: a review of literature," *Production and operations management*, vol. 12, no. 4, pp. 519–549, 2003.

[22] Centers for Disease Control and Prevention, "National hospital ambulatory medical care survey: 2011 emergency department summary tables. table 12." `http://www.cdc.gov/nchs/data/ahcd/nhamcs_emergency/2011_ed_web_tables.pdf`, 2011. Accessed: May 28, 2016.

[23] Centers for Disease Control and Prevention, "Asthma and schools." http://www.cdc.gov/HealthySchools/asthma/, 2015. Accessed: May 28, 2016.

[24] Centers for Disease Control and Prevention, "Asthmastats." http://www.cdc.gov/asthma/asthma_stats/aststatchild_missed_school_days.pdf, 2015. Accessed: May 28, 2016.

[25] Centers for Disease Control and Prevention, "Down syndrome cases at birth increased." http://www.cdc.gov/features/dsdownsyndrome/, last reviewed: May 15, 2012. Accessed: August 19, 2014.

[26] Centers for Disease Control and Prevention, "World down syndrome day." http://www.cdc.gov/ncbddd/birthdefects/features/DownSyndromeWorldDay-2012.html, last updated: February 26, 2016. Accessed: August 19, 2014.

[27] Centers for Disease Control and Prevention, "Asthma and schools." http://www.cdc.gov/HealthySchools/asthma/, last updated: June 17, 2015. Accessed: May 28, 2016.

[28] Centers for Disease Control and Prevention, "Asthmastats." http://www.cdc.gov/asthma/asthma_stats/aststatchild_missed_school_days.pdf, last updated: October 5, 2015. Accessed: May 28, 2016.

[29] Centers for Disease Control and Prevention. National Center for Health Statistics, "VitalStats." http://www.cdc.gov/nchs/vitalstats.htm. Accessed: April 2, 2013.

[30] Centers for Disease Control and Prevention. National Center for Health Statistics, "VitalStats." http://www.cdc.gov/nchs/vitalstats/births.htm. Accessed: March 30, 2016.

[31] Chan, C. W. and Green, L. V., "Improving access to healthcare: models of adaptive behavior," in *Handbook of Healthcare Operations Management*, pp. 1–18, Springer, 2013.

[32] Deo, S., Iravani, S., Jiang, T., Smilowitz, K., and Samuelson, S., "Improving health outcomes through better capacity allocation in a community-based chronic care model," *Operations Research*, vol. 61, no. 6, pp. 1277–1294, 2013.

[33] Deo, S., Rajaram, K., Rath, S., Karmarkar, U. S., and Goetz, M. B., "Planning for hiv screening, testing, and care at the veterans health administration," *Operations research*, vol. 63, no. 2, pp. 287–304, 2015.

[34] Emory Univeristy. School of Medicine. Department of Human Genetics, "Down syndrome cases at birth increased." https://genetics.emory.

edu/documents/resources/Emory_Human_Genetics_Amniocentesis.PDF. Accessed: August 19, 2014.

[35] Farina, A., LeShane, E. S., Lambert-Messerlian, G. M., Canick, J. A., Lee, T., Neveux, L. M., Palomaki, G. E., and Bianchi, D. W., "Evaluation of cell-free fetal dna as a second-trimester maternal serum marker of down syndrome pregnancy," *Clinical Chemistry*, vol. 49, no. 2, pp. 239–242, 2003.

[36] Fitzpatrick, A., Garcia, E., Serban, N., and Swann, J., "A study of the impact of geographic access on severe health outcomes for pediatric asthma." Working paper, Georgia Institute of Technology, 2014.

[37] Garcia, E., Serban, N., Swann, J., and Fitzpatrick, A., "The effect of geographic access on severe health outcomes for pediatric asthma," *Journal of Allergy and Clinical Immunology*, vol. 136, no. 3, pp. 610–618, 2015.

[38] Gebb, J. and Dar, P., "Should the first-trimester aneuploidy screen be maternal age adjusted? screening by absolute risk versus risk adjusted to maternal age," *Prenatal Diagnosis*, vol. 29, no. 3, pp. 245–247, 2009.

[39] Gekas, J., Durand, A., Bujold, E., Vallée, M., Forest, J.-C., Rousseau, F., and Reinharz, D., "Cost-effectiveness and accuracy of prenatal down syndrome screening strategies: should the combined test continue to be widely used?," *American Journal of Obstetrics and Gynecology*, vol. 204, no. 2, pp. 175.e1–e8, 2011.

[40] Gekas, J., Gagné, G., Bujold, E., Douillard, D., Forest, J.-C., Reinharz, D., and Rousseau, F., "Comparison of different strategies in prenatal screening for downs syndrome: cost effectiveness analysis of computer simulation," *BMJ*, vol. 338, pp. 453–456, 2009.

[41] Georgia Department of Education, "Student enrollment by grade level." https://oraapp.doe.k12.ga.us/ows-bin/owa/fte_pack_enrollgrade.entry_form, 2016. Accessed: February 10, 2016.

[42] Georgia Department of Education, "Student enrollment by grade level." https://oraapp.doe.k12.ga.us/ows-bin/owa/fte_pack_enrollgrade.entry_form, last updated: October 4, 2016. Accessed: February 10, 2016.

[43] Georgia Department of Public Health, "2014 asthma prevalence report." https://dph.georgia.gov/sites/dph.georgia.gov/files/2014\%20Asthma\%20Prevalence\%20Report.pdf, 2014. Accessed: May 28, 2016.

[44] Georgia Department of Public Health, "2015 georgia asthma program & data summary." https://dph.georgia.gov/sites/dph.georgia.gov/files/related_files/site_page/Asthma.pdf, 2015. Accessed: May 28, 2016.

[45] Georgia Department of Public Health, "2015 georgia data summary: Asthma in children." https://dph.georgia.gov/sites/dph.georgia.gov/files/Final\%20-\%202015\%20Child\%20Asthma\%20Data\%20Summary.pdf, 2015. Accessed: May 28, 2016.

[46] Griffin, J., Keskinocak, P., and Swann, J., "Allocating scarce healthcare resources in developing countries: a case for malaria prevention," in *Handbook of Healthcare Operations Management*, pp. 511–532, Springer, 2013.

[47] Gupta, D. and Denton, B., "Appointment scheduling in health care: challenges and opportunities," *IIE transactions*, vol. 40, no. 9, pp. 800–819, 2008.

[48] Gurobi Optimization, I., "Gurobi optimizer reference manual," 2016.

[49] Harris, R. A., Washington, A. E., Nease, R. F., and Kuppermann, M., "Cost utility of prenatal diagnosis and the risk-based threshold," *The Lancet*, vol. 363, no. 9405, pp. 276–282, 2004.

[50] Ji, Y., Kumar, S., and Sethi, S., "Needle exchange for controlling hiv spread under endogenous infectivity," *INFOR: Information Systems and Operational Research*, pp. 1–25, 2016.

[51] Johns Hopkins Medicine, "First trimester screening for down syndrome and trisomies 13 & 18." http://www.hopkinsmedicine.org/gynecology_obstetrics/pdfs/first_trimester_screening_patients.pdf. Accessed: August 19, 2014.

[52] Jones, C. A., Clement, L. T., Hanley-Lopez, J., Morphew, T., Kwong, K. Y. C., Lifson, F., Opas, L., and Guterman, J. J., "The breathmobile program: Structure, implementation, and evolution of a large-scale, urban, pediatric asthma disease management program," *Disease Management*, vol. 8, no. 4, pp. 205–222, 2005.

[53] Jones, C. A., Clement, L. T., Morphew, T., Kwong, K. Y. C., Hanley-Lopez, J., Lifson, F., Opas, L., and Guterman, J. J., "Achieving and maintaining asthma control in an urban pediatric disease management program: the breathmobile program," *Journal of Allergy and Clinical Immunology*, vol. 119, no. 6, pp. 1445–1453, 2007.

[54] Kucukyazici, B. and Verter, V., "Managing community-based care for chronic diseases: the quantitative approach," in *Operations Research and Health Care Policy*, pp. 71–90, Springer, 2013.

[55] Kuppermann, M., Feeny, D., Gates, E., Posner, S. F., Blumberg, B., and Washington, A. E., "Preferences of women facing a prenatal diagnostic choice: long-term outcomes matter most," *Prenatal Diagnosis*, vol. 19, no. 8, pp. 711–716, 1999.

[56] Kuppermann, M., Goldberg, J. D., Nease Jr, R. F., and Washington, A. E., "Who should be offered prenatal diagnosis? the 35-year-old question.," *American Journal of Public Health*, vol. 89, no. 2, pp. 160–163, 1999.

[57] Kuppermann, M., Nease, R. F., Learman, L. A., Gates, E., Blumberg, B., and Washington, A. E., "Procedure-related miscarriages and down syndrome–affected births: implications for prenatal testing based on womens preferences," *Obstetrics & Gynecology*, vol. 96, no. 4, pp. 511–516, 2000.

[58] Kuppermann, M., Nease Jr, R. F., Gates, E., Learman, L. A., Blumberg, B., Gildengorin, V., and Washington, A. E., "How do women of diverse backgrounds value prenatal testing outcomes?," *Prenatal Diagnosis*, vol. 24, no. 6, pp. 424–429, 2004.

[59] Kuppermann, M. and Norton, M. E., "Prenatal testing guidelines: time for a new approach," *Gynecologic and Obstetric Investigation*, vol. 60, no. 1, pp. 6–10, 2005.

[60] Mayo Clinic, "Childhood asthma." `http://www.mayoclinic.org/diseases-conditions/childhood-asthma/home/ovc-20193095`, 2016. Accessed: May 28, 2016.

[61] Morris, J., Wald, N., and Watt, H., "Fetal loss in down syndrome pregnancies," *Prenatal Diagnosis*, vol. 19, no. 2, pp. 142–145, 1999.

[62] Morris, J., Mutton, D., and Alberman, E., "Corrections to maternal age-specific live birth prevalence of Down's syndrome," *Journal of Medical Screening*, vol. 12, no. 4, pp. 202–202, 2005.

[63] National Center for Education Statistics, "Common Core of Data (CCD) 2014-15 school year." `https://nces.ed.gov/ccd/pubschuniv.asp`. Accessed: May 28, 2016.

[64] National Center for Health Statistics, "Asthma." `http://www.cdc.gov/nchs/fastats/asthma.htm`, 2016. Accessed: May 28, 2016.

[65] National Center for Health Statistics. Centers for Disease Control and Prevention, "Asthma." `http://www.cdc.gov/nchs/fastats/asthma.htm`, last updated: February 10, 2016. Accessed: May 28, 2016.

[66] Natoli, J. L., Ackerman, D. L., McDermott, S., and Edwards, J. G., "Prenatal diagnosis of down syndrome: a systematic review of termination rates (1995-2011)," *Prenatal Diagnosis*, vol. 32, no. 2, pp. 142–153, 2012.

[67] Odibo, A. O., Gray, D. L., Dicke, J. M., Stamilio, D. M., Macones, G. A., and Crane, J. P., "Revisiting the fetal loss rate after second-trimester genetic amniocentesis: A single centers 16-year experience," *Obstetrics & Gynecology*, vol. 111, no. 3, pp. 589–595, 2008.

[68] Ohno, M. and Caughey, A., "The role of noninvasive prenatal testing as a diagnostic versus a screening tool–a cost-effectiveness analysis," *Prenatal Diagnosis*, vol. 33, no. 7, pp. 630–635, 2013.

[69] Padula, F., Laganà, A. S., Vitale, S. G., D'Emidio, L., Coco, C., Giannarelli, D., Cariola, M., Favilli, A., and Giorlandino, C., "The introduction of the absolute risk for the detection of fetal aneuploidies in the first-trimester screening," *The Journal of Maternal-Fetal & Neonatal Medicine*, vol. 30, no. 10, pp. 1249–1253, 2017.

[70] Paltiel, A. D., Kuntz, K. M., Weiss, S. T., and Fuhlbrigge, A. L., "An asthma policy model," in *Operations Research and Health Care*, pp. 659–693, Springer, 2005.

[71] Parker, S. E., Mai, C. T., Canfield, M. A., Rickard, R., Wang, Y., Meyer, R. E., Anderson, P., Mason, C. A., Collins, J. S., Kirby, R. S., and Correa, A., "Updated National Birth Prevalence Estimates for Selected Birth Defects in the United States, 2004-2006," *Birth Defects Research Part A: Clinical and Molecular Teratology*, vol. 88, no. 12, pp. 1008–1016, 2010.

[72] Patel, B., Sheridan, P., Detjen, P., Donnersberger, D., Gluck, E., Malamut, K., Whyte, S., Miller, A., and Qing, H., "Success of a comprehensive school-based asthma intervention on clinical markers and resource utilization for inner-city children with asthma in chicago: the mobile care foundation's asthma management program," *Journal of Asthma*, vol. 44, no. 2, pp. 113–118, 2007.

[73] Presson, A. P., Partyka, G., Jensen, K. M., Devine, O. J., Rasmussen, S. A., McCabe, L. L., and McCabe, E. R., "Current Estimate of Down Syndrome Population Prevalence in the United States," *The Journal of Pediatrics*, vol. 163, no. 4, pp. 1163–1168, 2013.

[74] Puterman, M. L., *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 1994.

[75] Ronald McDonald House Charities, "Home page." `http://www.rmhc.org/home`, 2016. Accessed: May 28, 2016.

[76] Ronald McDonald House Charities, "Ronald mcdonald care mobile: Providing access to health care where children need it most." `http://www.rmhc.org/ronald-mcdonald-care-mobile`, 2016. Accessed: May 28, 2016.

[77] Ronald McDonald House Charities (Atlanta), "The ronald mcdonald care mobile is bringing asthma care to children where they live, learn, and play." `https://armhc.org/caremobile/`, last updated: 2017. Accessed: Aug 22, 2017.

[78] Ronald McDonald House Charities (Atlanta), "Care mobile unveiled." `https://armhc.org/blog/care-mobile-unveiled/`, last updated: November, 2016. Accessed: Aug 22, 2017.

[79] Ryzhov, I. O., Frazier, P. I., and Powell, W. B., "On the robustness of a one-period look-ahead policy in multi-armed bandit problems," *Procedia Computer Science*, vol. 1, no. 1, pp. 1635–1644, 2010.

[80] Ryzhov, I. O. and Powell, W., "The knowledge gradient algorithm for online subset selection," in *Adaptive Dynamic Programming and Reinforcement Learning, 2009. ADPRL'09. IEEE Symposium on*, pp. 137–144, IEEE, 2009.

[81] Ryzhov, I. O., Powell, W. B., and Frazier, P. I., "The knowledge gradient algorithm for a general class of online learning problems," *Operations Research*, vol. 60, no. 1, pp. 180–195, 2012.

[82] Savelsbergh, M. and Smilowitz, K., "Stratified patient appointment scheduling for mobile community-based chronic disease management programs," *IIE Transactions on Healthcare Systems Engineering*, vol. 6, no. 2, pp. 65–78, 2016.

[83] Schmidt, P., Hörmansdörfer, C., Golatta, M., and Scharf, A., "Analysis of the distribution shift of detected aneuploidies by age independent first trimester screening," *Archives of Gynecology and Bbstetrics*, vol. 281, no. 3, pp. 393–399, 2010.

[84] Song, K., Musci, T. J., and Caughey, A. B., "Clinical utility and cost of non-invasive prenatal testing with cfdna analysis in high-risk women based on a us population," *The Journal of Maternal-Fetal & Neonatal Medicine*, vol. 26, no. 12, pp. 1180–1185, 2013.

[85] Spencer, K., "Age related detection and false positive rates when screening for Down's Syndrome in the first trimester using fetal nuchal translucency and maternal serum free βhCG and PAPP-A," *BJOG: An International Journal of Obstetrics & Gynaecology*, vol. 108, no. 10, pp. 1043–1046, 2001.

[86] St. Louis Children's Hospital, "Healthy kids express - asthma program." `http://www.rmhc.org/ronald-mcdonald-care-mobile`, 2016. Accessed: May 28, 2016.

[87] United States Department of Health and Human Services (US DHHS), Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS), Division of Vital Statistics, "Natality public-use data 2014, on CDC WONDER Online Database." `http://wonder.cdc.gov/natality-current.html`. Accessed: July 4, 2017.

[88] Wald, N. and Hackshaw, A., "Tests using multiple markers," in *Antenatal and Neonatal Screening* (Wald, N. J. and Leck, I., eds.), pp. 23–57, Oxford: Oxford University Press, 2000.

[89] Wald, N. J., Rudnicka, A. R., and Bestwick, J. P., "Sequential and contingent prenatal screening for down syndrome," *Prenatal Diagnosis*, vol. 26, no. 9, pp. 769–777, 2006.

[90] Wald, N., Rodeck, C., Hackshaw, A., Walters, J., Chitty, L., and Mackinson, A., "First and second trimester antenatal screening for down's syndrome: the results of the serum, urine and ultrasound screening study (suruss)," *Journal of Medical Screening*, vol. 10, no. 2, pp. 55–104, 2003.

[91] Wald, N., Rodeck, C., Hackshaw, A., Walters, J., Chitty, L., and Mackinson, A., "First and second trimester antenatal screening for down's syndrome: the results of the serum, urine and ultrasound screening study (suruss)," *Health Technology Assessment*, vol. 7, no. 11, pp. 1–88, 2003.

[92] Walker, B., Ashwood, E. R., Jackson, B. R., and Lagrave, D., "A tradeoff analysis of risk cutoffs for the quadruple serum screen for down syndrome," *Prenatal Diagnosis*, vol. 33, no. 12, pp. 1201–1206, 2013.

[93] Yan, J., Ayer, T., Keskinocak, P., and Caughey, A. B., "Preference-sensitive risk-cutoff values for prenatal-integrated screening test for Down syndrome," *Prenatal Diagnosis*, vol. 35, no. 7, pp. 645–651, 2015.

[94] Yan, J., Ayer, T., Keskinocak, P., and Caughey, A. B., "Age-based differences in the predictive accuracy of a one-size-fits-all risk-cutoff value in prenatal integrated screening for Down syndrome," *Prenatal Diagnosis*, p. to appear, 2017.

[95] Zhang, Y., Wu, H., Denton, B. T., Wilson, J. R., and Lobo, J. M., "Conducting probabilistic sensitivity analysis for decision models based on markov chains," 2015.

VITA

Jia Yan was born and grew up in China. She received her Bachelor and Master degrees in Industrial Engineering from Tsinghua University (Beijing, China) in 2009 and 2011, respectively. After that she came to the United States and began to pursue her Ph.D. in Operations Research in the H. Milton Stewart School of Industrial and Systems Engineering at Georgia Institute of Technology. In 2014, she received a Master degree in Operations Research from Georgia Institute of Technology. She won 2017 Graduate Student Paper Competition of Society for Health Systems with the paper "Age-Specific Risk-Cutoff Values in Prenatal-Integrated Screening for Down Syndrome Considering Fairness". She is interested in learning and applying operations research and statistical methodologies to various areas and making impacts.